THREE ESSAYS IN FINANCIAL ECONOMICS AND LAW

By

WEI-LIN LIU

A DISSERTATION PRESENTED TO THE GRADUATE SCHOOL
OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

1997

To My Wife Yan Li

ACKNOWLEDGMENTS

I am deeply indebted to Tracy Lewis and Charles Hadlock, who provided guidance, support and inspiration. My gratitude to both of them for their generous help throughout this work is beyond what words can describe. I also wish to express my thanks to David Sappington, Joel Houston and members of my committee.

TABLE OF CONTENTS

ACKNOWLEDGMENTS....iii

ABSTRACT....vi

CHAPTERS

1 INTRODUCTION....1
Motivating and Compensation Investment Advisors....1
Setting Standards for Credible Compliance and Law Enforcement....2
Monitoring and The Optimal Mix of Public and Private Debt Claims....3

2 MOTIVATING AND COMPENSATING INVESTMENT ADVISORS....5
Introduction....5
A Simple Discrete Model....7
The General Two-State Model....13
Extension....24
Conclusions....26

3 SETTING STANDARDS FOR CREDIBLE COMPLIANCE AND LAW ENFORCEMENT....34
Introduction....34
Elements of The Basic Model....37
Analysis of The Simple Case....42
Privately Informed Enforcer....46
Heterogenous Parties....49
Setting Optimal Fines....53
Conclusions....55

4 MONITORING AND THE OPTIMAL MIX OF PUBLIC AND PRIVATE DEBT CLAIMS....58
Introduction....58
Elements of The Model....64
The Mix of Long-Term Public and Bank Debt Claims....69
The Optimal Bank Debt....73

The Optimal Mix....85
Empirical Implications....90
Conclusions....92

5 CONCLUSIONS....96

APPENDIX A: PROOFS O F THE MAIN RESULTS IN CHAPTER TWO....98

APPENDIX B: PROOFS O F THE MAIN RESULTS IN CHAPTER THREE....108

APPENDIX C: PROOFS O F THE MAIN RESULTS IN CHAPTER FOUR....117

REFERENCES....125

BIOGRAPHICAL SKETCH....129

Abstract of Dissertation Presented to the Graduate School of
the University of Florida in Partial Fulfillment of the
Requirements for the Degree of Doctor of Philosophy

THREE ESSAYS IN FINANCIAL ECONOMICS AND LAW

By

WEI-LIN LIU

August 1997

Chairman: David Brown
Co-Chairman: Tracy Lewis
Major Department: Finance

The dissertation is a collection of studies in financial economics and law. The first chapter introduces the thesis. The second chapter analyzes the problem of how an investor can compensate an investment advisor to both motivate the advisor to diligently collect information and elicit truthful revelation of his private information. I find that the structure of an optimal compensation scheme depends both on the technology of information collection and on the accuracy level of the information that the advisor is induced to achieve. I identify instances in which the design of an optimal compensation scheme is independent of whether the information acquired by the advisor is publicly observable or not.

The third chapter examines the problem of setting optimal legal standards when enforcers of the standards must be motivated to detect violations. I find that some divergence between the marginal benefits and the marginal costs of providing care by potential violators is required to control enforcement costs. Furthermore, the setting of

standards may effectively substitute for the setting of fines when penalties for violations are fixed. In particular, it is found that imposing maximal fines may be welfare reducing.

The fourth chapter explores the issue of how firms optimally design their debt structures by using both public and private debt. The principal finding is that in general the optimal debt structures are a mix of long-term public debt and private debt with varying repayment schedules. In addition, I show how financial intermediaries can produce information through monitoring the firms they lend to, and I extend the notion of delegated monitoring to the case when there are multiple classes of creditors.

CHAPTER 1

INTRODUCTION

This dissertation is a collection of studies in financial economics and law. The second chapter analyzes the problem of how best an investor can compensate an investment advisor to both motivate the advisor to diligently collect information and elicit truthful revelation of his private information. The third chapter examines the problem of setting legal standards when enforcers of the standards must be motivated to detect violations. The fourth chapter explores the issue of how firms optimally design their debt structures by using both public and private debt.

Motivating and Compensating Investment Advisors

Individual investors generally do not have the expertise to assess prospective investment opportunities, and they may rely on investment advisors for expert opinions. In designing the advisor's compensation scheme, however, two problems arise: 1) it must motivate the advisor to diligently collect information, and 2) it must induce the advisor to truthfully reveal his information to the investor. In chapter 2, I derive the optimal compensations schemes which satisfy both of these conditions.

I first analyze the case in which the advisor's information is publicly observable. I find that the optimal payment scheme rewards the advisor more richly for correctly predicting an outcome, if expending effort best enhances his ability to predict that outcome.

When the advisor's information is not publicly observable, I find that the need to induce the advisor to expend effort generally interferes with the need to elicit truthful revelation. I show that in general there exists a critical level of effort. If the advisor is induced to expend an effort higher than the critical one, whether his information is publicly observable or not is immaterial. If the advisor is induced to expend an effort lower than the critical one, the two needs interact. In this case, the advisor is rewarded more richly if he correctly predicts the outcome less likely to occur.

The analysis in chapter 2 extends previous analysis by Kilhstrom (1986) who assumes that the advisor's information is publicly observable and that the advisor's effort improves the accuracy of his information equally. The results in Bhattacharya and Pfleiderer (1987) and Stoughton (1995) rely critically on the symmetry property of the information technology. In my analysis, this assumption is relaxed in a binary signal setting.

Setting Standards for Credible Compliance and Law Enforcement

Chapter 3 analyzes the problem of setting socially optimal legal standards when enforcers of the standards must be motivated to oversee potential violators. Beginning with Becker (1968), research in enforcement of legal standards has focused on the setting of fines as a primary tool of enforcement. In contrast, this analysis characterizes how the setting of legal standards affects the behavior of complying parties, law enforcers, and the net social surplus generated by the regulations.

The analysis in chapter 3 reveals that it is desirable to induce potential violators of the standards to provide care levels that either exceed or fall short of the surplus maximizing level. In some instances, a slight loosening of standards will decrease enforcement costs.

Such instances arise whenever looser standards cause enforcers to reduce their effort because the marginal returns from monitoring decrease as the probability of noncompliance decreases. For other applications, monitoring effort may fall as the probability of noncompliance increases. For these cases, it will be desirable to set tighter standards and induce greater care in order to reduce the enforcers' expenditures on effort.

Two extensions of this result are also presented. In the first instance, the possibility that the costs of monitoring effort vary by the enforcers' abilities to observe and process information is considered. These costs are known privately by the enforcers. It is shown that the presence of asymmetric information reinforces the main finding that standards are distorted to reduce enforcement costs. In the second extension, the possibility that parties differ in the costs they incur in taking care is examined. It is shown that the main finding can be generalized to this case as well.

In addition, the analysis examines the relationship between fines and standards. It is shown that imposing the largest fine is not necessarily desirable because increases in fines may increase costly enforcement effort.

Monitoring and the Optimal Mix of Public and Private Debt Claims

Chapter 4 explores the issue of how firms optimally design their debt structures by using both public debt and private debt. The theory of finance suggests that financial intermediaries can act as delegated monitors and provide flexibility allowing for modifications of loan contracts when needed. However, previous studies have only provided partial analysis of private lenders' incentives to monitor. In addition, existing literature does not provide adequate explanations of the need for firms to finance with both private debt and

public debt. Moreover, no explanation has been offered to justify the notion of delegated monitoring in the presence of multiple classes of debtors.

This analysis provides some answers to these important issues. My analysis is based on the assumption that firms' objective in designing optimal debt structures is to credibly commit to repay debtors at minimum cost. I find that the optimal debt structure is in general strict mixes of public and private debt. Public debt provides several benefits: 1), it allows the firm to pay out profits without perturbing bank's incentive for monitoring; 2) it allows the firm to maximize the lenders' total benefit per unit of monitoring effort, thereby reducing the amount of monitoring required to ensure initial financing. I show that in the optimal debt structure, the public debt is generally long-term while the maturity of private debt depends on the severity of agency problems. In addition, I find that when firms raise external financing by optimally combining public and private debt, they prefer to align the public lenders incentive over monitoring with that of the private lenders. Thus, private lenders act as delegated monitors in the presence of multiple classes of creditors.

My analysis is closely related to that of Hart and Moore (1991). Besides sharing similar assumptions, my analysis extends Hart and Moore's analysis by accommodating asymmetric information and by recognizing the negotiation costs involved in dealing with dispersed public debt holders. My analysis also differs from those of Park(1994) and Raja and Winton (1995) in that I emphasize the inherent incompleteness of debt contracts and do not rely on covenants to provide banks with monitoring incentives.

CHAPTER 2
MOTIVATING AND COMPENSATING INVESTMENT ADVISORS

Introduction

Investing profitably requires accurate information. Often, an investor may not have the knowledge or the skill to collect and process relevant information about investment opportunities. The investor must rely on the expertise of an investment advisor. In structuring optimal compensation schemes for the advisor, two problems arise: 1) inducing the advisor to diligently collect information; and 2) inducing the advisor to reveal his information to the investor. This paper characterizes optimal compensation schemes which satisfy both of these two requirements.

To characterize the optimal compensation schemes, we employ the following model. An investor can invest in a risk free asset with a known return and a risky asset whose return depends on the states of nature. Initially, the investor and the advisor share the same prior about the probability of each state occurring. By expending effort, however, the advisor can privately observe a signal correlated with the state of nature. Expending greater effort improves the accuracy of the signal. The investor makes his investment decisions based on the advisor's reported signal. Subsequently, the actual return from the risky asset is publicly observed, and the advisor is compensated based upon his report and the realized return[1].

[1]Alternatively, our model can be interpreted as one in which no communication takes place. The investor pre-announces the investment schedule and the advisor implements the schedule after a signal is acquired. The decision made by the advisor in implementing

When the signal is publicly observable, the investor's only concern is to motivate the advisor to expend effort. In this case, we find that the advisor is more richly rewarded for correctly predicting a state if exerting effort produces the greatest percentage increase in his ability to predict that state. This follows because to best motivate the advisor, the investor desires to compensate him more for a correct prediction which is most indicative of the effort expended. Thus, the optimal compensation scheme rewards the advisor more for a correct prediction if the probability of achieving it is most sensitive to the effort expended.

When the signal is not publicly observable, the investor must be concerned with both motivating the advisor and eliciting truthful revelation of the private signal. We find that there exists a critical effort level: When the advisor is induced to expend an effort higher than the critical one, the investor's inability to observe the signal is inconsequential for the design of optimal payment scheme; In contrast, when the advisor is induced to expend an effort lower than the critical one, the need to elicit truthful revelation of the observed signal interacts with the need to motivate the advisor. In the latter case, the payment to the advisor depends both on whether he predicts the state correctly and on which state occurs. The payment is larger when the advisor correctly predicts the less likely state. The intuition for this result is the following. If the advisor is rewarded a fixed amount whenever he makes a correct prediction, he will choose to predict the state most likely to occur (based on the prior) if he shirks. In doing so, however, the advisor loses the opportunity of being rewarded for correctly predicting the less likely state. To motivate the advisor, the investor must impose sufficient opportunity cost on him for not expending effort. By rewarding the

---

the investment schedule is, however, observable and can be contracted upon.

advisor more for correctly predicting the less likely state, shirking and making predictions based only on the prior become less attractive. The new payment scheme can both motivate the advisor and induce truthful revelation.

Before proceeding, we relate our analysis to earlier studies in the literature. The study by Kilhstrom(1986), which is most closely related to our analysis, analyzes how to induce an advisor to work diligently when his informative signal is publicly observable. We extend his analysis by allowing the advisor's signal to be private information. The extension enables us to investigate how the need to elicit truthful revelation interacts with the need to motivate the advisor.

Also related to our analysis is the work of Bhattacharya and Pfleiderer(1985) who study the problem of screening of agents (advisors) endowed with information technologies that differ in the accuracy levels of the signals produced, and subsequently eliciting truthful revelation of the privately observed signals. Acquiring the signal is assumed to be costless. This model is modified in a later paper by Stoughton (1993) who studies a moral hazard problem similar to ours'. The results in both of the two studies, however, rely critically on the symmetric information structure that they employ, whereby the distribution of states of the nature is symmetric conditional on any signal. Our analysis, on the other hand, does not require this assumption.

A Simple Discrete Model

A risk neutral investor can invest in a risky asset or a riskless asset. The riskless asset yields a return R with certainty. The return of the risky asset depends on the state of nature: It returns $r_H > R$ in the favorable state and $r_L < R$ in the unfavorable state. The

investor's prior belief of the probability of the two states is π_H and π_L, respectively. For simplicity, we assume $\pi_H r_H + \pi_L r_L = R$.

The investor can acquire additional information about the state of nature by hiring an advisor. By expending effort, the advisor, who initially shares the same prior as the investor, can observe a signal x, with $x \in \{x_H, x_L\}$, correlated with the realized state. The correlation between the signal and the state is characterized by the conditional distribution function $f(x_i|r_j,e)$, $i,j \in \{H,L\}$. There are two effort levels, $e=e_L$ and $e=e_H>e_L$. For notational ease, we set $f(x_i|r_i,e_H)=p_i$ and $f(x_j|r_j,e_L)=q_j$, $i,j \in \{H,L\}$. The advisor's effort improves the accuracy of the signal in the Blackwell sense. We capture this idea by assuming $p_i>q_i$, $i \in \{H,L\}$.

The advisor's utility function $V_A(W_A)-C(e)$ is separable in income and effort, where $V_A(W_A)$ is the utility of the end of period payoff W_A and C(e) is the cost for expending effort e. We assume that the advisor is liquidity constrained and enjoys limited liability protection[2]. The advisor's reservation utility is normalized to zero.

The investor can not observe the advisor's effort and the signal he observes. He makes the investment decision based on the advisor report. Contingent upon a report x, he chooses to invest $\lambda(x)$ in the risky asset. The investor has an endowment of $W_0>0$, and borrowing and short-selling are not allowed, so that $0 \le \lambda(x) \le W_0$.

The timing in the model is as follows. First, the investor makes a take-it-or-leave-it offer to the advisor, specifying a compensation scheme w(x,r), which depends on both the

---

[2]This assumption implies that the advisor can not signal the investor his private information by taking a position in the risky asset. For the latter approach, see Leland and Pyle (1977), Allen (1990).

advisor's report and the state of nature publicly observed ex post. Second, if the advisor accepts the offer, he selects the level of effort to expend. Otherwise, the game is terminated. Third, the advisor observes a signal and makes a report to the investor. The advisor's report amounts to a prediction of future state. Reporting x_H (x_L) amounts to predicting the occurrence of good (bad) state. Fourth, the investor makes and implements the optimal investment decision based on the advisor's report. Finally, the realized return of the risky asset is publicly observed and the advisor is paid as promised.

The optimal investment decision is determined, in equilibrium, for a given equilibrium strategy of the advisor . Specifically, the advisor is assumed to truthfully reveal his privately observed signal. Recall, a risk neutral investor ranks the investment opportunities by their expected payoffs. This implies that the optimal investment decision is: $\lambda(x_L)=0$ and $\lambda(x_H)=W_0$, which is independent of the advisor's compensations.

We assume that it is desirable to motivate the advisor to acquire the more informative signals. The investor's objective is to induce the advisor to expend effort and truthfully reveal his signals at minimum cost. Inducing truthful revelation implies the following constraint.

$$E_r[w(x,r)|e_H,x] \geq E_r[w(x',r)|e_H,x] \qquad \forall x,x'. \tag{2-1}$$

Inducing the advisor to expend effort requires

$$E_{x,r}[w(x,r)|e_H,x]-E \geq E_x\{Max_{x'}E_r[w(x',r)|x,e_L]|e_L\} \tag{2-2}$$

Finally, we add the individual rationality constraint, which ensures that the advisor will accept the contract, and the limited liability constraint.

$$E_{x,r}[w(x,r)|e_H,x]-E \geq 0 \tag{2-3}$$

$$w(x,r) \geq 0, \qquad \forall x,r \tag{2-4}$$

The investor's problem is then

$$\text{Min } E_{x,r}[w(x,r)]$$

s.t (2-1), (2-2), (2-3) and (2-4).

Lemma 2.1: Constraint (2-1) implies $w(x_H,r_H) \geq w(x_L,r_H)$ and $w(x_L,r_L) \geq w(x_H,r_L)$

The proof of lemma 1 is straightforward, and is omitted. The (2-4) constraint implies that the right hand side of (2-2) is nonnegative. Thus, (2-2) and (2-4) imply (2-3), and we will ignore (2-3). Simple rearrangement indicates both (2-1) and (2-2) can be expressed in terms of the two differences $w(x_H,r_H)-w(x_L,r_H)$ and $w(x_L,r_L)-w(x_H,r_L)$. This implies that (2-2) and (2-1) continue to hold under a simultaneous decrease of $w(x,r)$, $\forall x,r$, provided the two differences remain unchanged. Lemma 2.1 and investor's cost minimizing then imply $w(x_L,r_H)=w(x_L,r_H)=0$ in the optimal contract. That is, the advisor will not be compensated if he makes a wrong prediction. The advisor is only compensated when he makes a correct prediction. In the following, we denote $w(x_H,r_H)$ by Y_G and $w(x_L,r_L)$ by Y_B. Y_G and Y_B represent the bonuses awarded to the advisor for correctly predicting the good and the bad state respectively.

Constraint (2-1) implies the following two inequalities,

$$p_H \pi_H Y_G \geq (1-p_L)\pi_L Y_B,$$

$$p_L \pi_L Y_B \geq (1-p_H)\pi_H Y_G.$$

These two constraints ensure that the advisor will truthfully reveal signals x_H and x_L given he has expended effort e_H. Contracts satisfying the two inequalities are represented by the shaded area in figure 2-1. Contracts satisfying the first constraint lie under the line OH.

Contracts satisfying the second lie above OL. The small cone enclosed by the dotted lines OH′ and OL′ contains contracts that induce truthful revelation at effort e_L[3]. Figure 2-1 reveals that the set of truth-telling contracts at e_L is a subset of the set of truth-telling contracts at e_H. This implies that it is easier to induce truthful revelation when the advisor is better informed. Figure 2-1 also indicates that simple profit sharing contracts are not incentive compatible. Consider, the payment scheme corresponding to point K. This payment scheme awards the advisor only when investment yields a net profit. Such a contract does not induce the advisor to truthfully reveal an unfavorable signal and results in over-investment in the risky asset.

Feasible contracts must satisfy constraint (2-2) in addition to (2-1). Consider first contracts lie in the cone enclosed by the dotted lines in fig.2-1. These contracts induce truthful revelation at both the high and the low effort levels. For these contracts, constraint (2-2) implies the following inequality,

$$(p_H-q_H)\pi_H Y_G+(p_L-q_L)\pi_L Y_B \geq E$$

Contracts satisfying this inequality lie above the line AB (figure 2-2). Thus, the first subset of feasible contracts lie inside the inner cone and above the line AB. Next, consider contracts lie above OH′ but under OH. These contracts induce truth-telling at e_H but not at e_L. At e_L, the advisor always prefers to report x_L. For these contracts, (2-2) implies

$$[p_H\pi_H Y_G-(1-q_L)\pi_L Y_B]+(p_L-q_L)\pi_L Y_B \geq E$$

Contracts satisfying this constraint lie under line AC. Thus, the second part of feasible contracts lie above AH′ and under AC. Similarly, contracts lying under OL′ but above OL

---

[3]These contracts satisfy the same constraint as (2-1) with p replaced by q.

induce truthful revelation only at e_H. At e_L, the advisor always prefers to report x_H. Applying (2-2) to these contracts provides

$$(p_H-q_H)\pi_H Y_G+[p_L\pi_L Y_B-(1-q_L)\pi_H Y_G]\geq E$$

The final part of feasible contract lies under BL' and above BD. Feasible contracts which satisfy all constraints are represented by the shaded area in figure 2-2. This shaded area lies strictly inside the cone which contains the contracts eliciting truthful revelation at effort e_H. It follows that constraint (2-2) implies (2-1). A formal proof of this is quite simple and is omitted.

The investor's problem is to find the minimum cost contract from contracts in the shaded region in figure 2-2. Figure 2-3 depicts the investor's iso-cost lines. The cost associated with the iso-cost line increases in the north-east direction. It is instructive to consider first the case in which the advisor's signal is publicly observable. In this case, the investor's only concern is to induce the advisor to choose the high effort level. All payment schemes lying above the line MN are feasible. It is evident from figure 2-3, the optimal contract corresponds to either point M or point N depending on the slopes of iso-cost lines. If the iso-cost lines are steeper than the line MN, point M represents the optimal payment scheme. Such a contract compensates the advisor only when he correctly predicts the unfavorable state. Simple calculation reveals that iso-cost lines are steeper if $p_H/p_L\geq(p_H-q_H)/(p_L-q_L)$. This condition can be written in a more suggestive form, $(p_L-q_L)/p_L\geq(p_H-q_H)/p_H$. The left hand side of the inequality is the percentage improvement of the accuracy of signal x_L. The right hand side is the percentage improvement of the accuracy of signal x_H. Thus, the advisor is more richly rewarded for correctly predicting bad investments if exerting extra

effort produces the greatest percentage increase in the advisor's ability to predict the bad state. Correctly predicting the bad state is more indicative that the advisor is diligently collecting information. Similarly, if exerting high level of effort produces the greatest percentage increase in the advisor's ability to predict the good state, he is more richly rewarded when he correctly predicts the good state. In figure 2-3, this corresponds to the flatter iso-cost lines. Returning to the case when the advisor's signal is not publicly observable, feasible contracts lie within the shaded cone. Since points M and N lie outside of the shaded area, the optimal payment scheme in the previous case is no longer incentive compatible. Instead, the optimal payment scheme is the point at which the iso-cost line is just touching the shaded cone. Again, depending on the slope of iso-cost lines, the optimal contract corresponds to either A or B. Figure 2-3 indicates that the investor continues to reward the advisor more richly for correctly predicting a state if exerting high level of effort produces the greatest percentage increase in the advisor's ability to predict that state. In contrast to the previous case, however, the advisor is compensated for both correctly predicting the favorable and the unfavorable state. Thus, the need to induce the advisor to expend effort interacts with the need to elicit truthful revelation. The interaction makes it more costly to induce the advisor to expend effort (as indicated in figure 2-3, the new optimal contract lies on a higher iso-cost line).

The General Two State Model

In this section, we extend our analysis to a more general setting. Specifically, we assume there is a continuum of effort level, $e \in [0,+\infty)$. To focus on the interaction between the two incentive concerns, we assume the correlation between the signal and the states

satisfies $f(x_L|r_L,e)=f(x_H|r_H,e)=\theta(e)$ with $\theta'(e)>0$, $\theta''(e)<0$ and $\lim_{e\to+\infty}\theta(e)=1$. Thus, expending effort improves the accuracy of the two signals equally. The concavity condition reflects decreasing marginal returns to effort. The advisor is strictly risk averse, with $V_A'(W_A)>0$ and $V_A''(W_A)<0$. The cost function C(e) satisfies $C''(e)>0$, $C'''(e)\geq 0$ and $C'(e)>0$ (except for e=0). Furthermore, we assume that $V_A(0)=C(0)=C'(0)=0$ and $V_A'(0)=+\infty$[4]. It will be convenient to regard the cost as a function of θ instead of e. The one-to-one correspondence between e and θ ensures that the function $C(\theta)$ is well defined. It follows from our assumption that $C(\theta)$ satisfies $C'(\theta)>0$ (except for $\theta=1/2$), $C''(\theta)>0$ and $C(\frac{1}{2})=C'(\frac{1}{2})=0$. The inverse of V_A will be denoted by h, and we assume that h is thrice continuously differentiable. We will denote $W_0[\pi_H(r_H-R)+\pi_L(R-r_L)]$ by β. β is the marginal return from improving the accuracy of the signal. The investor's problem, [I-P], is formally stated as follows.

$$Max_{\theta,\lambda(\cdot),w(\cdot,\cdot)} E_{x,r}[W_0R+\lambda(x)(r-R)-w(x,r)|\theta]$$

subject to

$$x\in Argmax_{x'} E_r[V_A(w(x',r))|x,\theta] \qquad \forall x\in\{x_L,x_H\} \tag{2-5}$$

$$\theta\in Argmax_{\theta'} E_x\{Max_{x'} E_r[V_A(w(x',r))-C(\theta')|x,\theta']|\theta\} \tag{2-6}$$

$$E_{x,r}[V_A(w(x,r))-C(\theta)|\theta]\geq U=0 \tag{2-7}$$

$$w(x,r)\geq 0 \qquad x\in\{x_L,x_H\} \qquad r\in\{r_L,r_H\} \tag{2-8}$$

Before proceeding, we present two results. First, we show that the investor's problem

---

[4]The assumptions $C'(0)=0$, $V_A'(0)=+\infty$ and $\pi_H r_H+\pi_L r_L=R$ ensure that it is desirable for the investor to hire the advisor.

has the following equivalent formulation [I-P']:

$$Max_{\theta,\lambda(x),w(\cdot,\cdot)} E_{x,r}[W_0 R+\lambda(x)(r-R)-w(x,r)|\theta]$$

subject to

$$\Pi(x_H,x_L)\geq\Pi(x_i,x_j) \qquad \forall x_i,x_j\in\{x_L,x_H\} \tag{2-9}$$

$$\theta\in ArgMax_{\theta'}[L(x_H,x_L|\theta')-C(\theta')] \tag{2-10}$$

$$L(x_H,x_L|\theta)-C(\theta)\geq 0 \tag{2-11}$$

$$w(x,r)\geq 0 \qquad x\in\{x_L,x_H\} \qquad r\in[r_L,r_H] \tag{2-12}$$

where $$\Pi(x_i,x_j)=Max_{\theta'}[L(x_i,x_j|\theta')-C(\theta')] \qquad \forall x_i,x_j\in\{x_L,x_H\} \tag{2-13}$$

$$L(x_i,x_j|\theta')=f(x_H|\theta')E_r(V_A(w(x_i,r))|x_H,\theta')+f(x_L|\theta')E_r(V_A(w(x_j,r))|x_L,\theta') \quad \theta'\in[1/2,1)^5 \tag{2-14}$$

Proposition 2.1: The investor's problem [I-P] is equivalent to [I-P']⁶.

The advisor's strategy consists of choosing a level of accuracy in the first stage and subsequently choosing a reporting rule in the second stage. Let $\{\theta,(x_i,x_j)\}$ denote such a strategy, with $x_i,x_j\in\{x_H,x_L\}$. Under this strategy, the advisor chooses an accuracy level θ in the first stage and reports x_i or x_j when the acquired signal is x_H or x_L. It is easy to see that $L(x_i,x_j|\theta)$-$C(\theta)$ is the expected payoff to the advisor when he adopts the strategy $\{\theta,(x_i,x_j)\}$. On the other hand, $\Pi(x_i,x_j)$ is the maximum expected payoff to the advisor if he chooses a second stage rule of report (x_i,x_j), independent of the choice of θ in the first stage. Constraint (2-9) requires that the payment scheme $\Pi(x_i,x_j)$, as a function of the rule of report,

---

[5] $f(x_H|\theta')$ and $f(x_L|\theta')$ are the marginal probability of the occurrence of signal x_H and x_L at accuracy level θ'.

[6] Throughout the dissertation, all proofs are relegated to the appendix.

attains its maximum when the advisor reports truthfully. Constraint (2-10) requires that the advisor's payoff, when he always reports truthfully, attains a maximum at θ--the accuracy level induced by the investor's contract.

Second, we show the following lemma which will help to simplify the constraints.

Lemma 2.2:(1)$\Pi(x_L,x_L)=L(x_L,x_L|\theta=1/2)$ and $\Pi(x_H,x_H)=L(x_H,x_H|\theta=1/2)$;

(2)If a payment scheme satisfies constraints (2), $\Pi(x_H,x_L)\geq\Pi(x_H,x_H)$ and $\Pi(x_H,x_L)\geq\Pi(x_L,x_L)$ for $\theta>1/2$, then

i) $w(x_H,r_H)>w(x_L,r_H)$ and $w(x_L,r_L)>w(x_H,r_L)$;

ii)$\Pi(x_L,x_H)=L(x_L,x_H|\theta=1/2)$.

If the advisor always announces x_L or x_H independent of the actual signal observed, his expected profit must be independent of the accuracy of the signal. Since improving accuracy is costly, the advisor optimally exerts no effort. This is reflected in Lemma 2.2 (1).

Part (2) ii) of the lemma implies that the investor must pay the advisor a strictly positive bonus when he correctly predicts the state. This follows because the state r_H (r_L) is more likely to occur conditional on x_H (x_L). The bonus is necessary to induce the advisor to truthfully reveal his private signal[7].

Given Lemma 2.2, we can simplify [I-P'] to the following problem [I-P''].

$$Max_{\theta,\lambda(x),w(\cdot,\cdot)}E_{x,r}[W_0R+\lambda(x)(r-R)-w(x,r)|\theta]$$

subject to

---

[7]The ordering in the payment scheme is similar to the monotonicity observed in the standard models with adverse selection Baron and Myerson (1982), Laffont and Tirole (1986). The single crossing property there corresponds to the condition $\theta>1/2$.

$$E_{x,r}[V_A(w(x,r))|\theta]-C(\theta)\geq E_x\{Max_{x'}E_r[V_A(w(x',r))|x,\theta=\frac{1}{2}]|\theta=\frac{1}{2}\} \qquad (2\text{-}15)$$

$$\theta\in Arg\max_{\theta'}E_{x,r}[V_A(w(x,r))|\theta'] \qquad (2\text{-}16)$$

$$w(x,r)\geq 0 \qquad x\in\{x_L,x_H\} \qquad r\in\{r_L,r_H\} \qquad (2\text{-}17)$$

In simplifying [I-P′], we have dropped the individual rationality constraint (2-11) since it is guaranteed by the limited liability constraint. Constraints in [I-P″] closely resemble those in the discrete model. The effort levels e_H and e_L corresponds to the effort level which generates the investor's desired accuracy and zero effort level. The difference arises from the continuous nature of effort levels. Again, it is instructive to consider the problem in which the advisor's signal is publicly observable--reduced problem. The reduced problem, [RP], is similar to [I-P″] but without the first constraint. Without the limited liability constraint, the reduced problem corresponds exactly to the problem investigated by Kilhstrom (1986). The solution of the reduced problem is summarized in the following lemma. In the lemma, θ_{RP} refers to the optimal accuracy level.

Lemma 2.3: At the solution to the reduced problem:

i)the compensation scheme is $w(x_H,r_L)=w(x_L,r_H)=0$ and $w(x_L,r_L)=w(x_H,r_H)=h(C'(\theta_{RP}))$;

ii)the optimal level of accuracy θ_{RP} solves $Max_\theta[\theta\beta-\theta h(C'(\theta))]$ and $1/2<\theta_{RP}<\theta_{FB}<1$.

Proof of the lemma follows from Kilhstrom (1984). However, for completeness, we provide a derivation in the appendix.

Based on our analysis in the discrete model, the result in lemma 2.3 is well anticipated. Since expending effort improves the accuracy of the two signals equally, our previous analysis indicates that the investor should be indifferent between rewarding the

advisor for correctly predicting the favorable or the unfavorable state. Advisor's risk aversion implies that he should be subject to minimum risk exposure. Therefore, the investor pays the advisor a fixed amount whenever he makes a correct prediction, i.e $w(x_L,r_L)=w(x_H,r_H)$. The result of lemma 2.3 is illustrated in figure 2-4.

To solve [I-P″], we adopt the two step approach in Grossman and Hart (1983). First we characterize the optimal payment scheme for implementing a given level of accuracy. Second we determine the optimal accuracy level. The problem for the first step is to minimize $E_{x,r}[w(x,r)|\theta]$ subject to the same constraints as in [I-P″].

Since the solution in the case $\pi_L \geq \pi_H$ is exactly symmetric to that in the case $\pi_H \geq \pi_L$[8], we assume $\pi_H \geq \pi_L$ in the ensuing analysis. Simple rearrangements reveal that all the constraints can be expressed in terms of $V_A(w(x_H,r_H))-V_A(w(x_L,r_H))$ and $V_A(w(x_L,r_L))$ $-V_A(w(x_H,r_L))$. This implies that the payment scheme remains to be feasible under a simultaneous decrease of $w(x,r)\ \forall x,r$, provided the two differences remain unchanged. The following result is then obvious.

Lemma 2.4.1: In the optimal solution, $w(x_H,r_L)=w(x_L,r_H)=0$.

Similar to the discrete model, we denote $V_A(w(x_H,r_H))$ by Y_G' and $V_A(w(x_L,r_L))$ by Y_B'. Y_G' and Y_B' represent the two bonuses received when the advisor correctly predicts the two states. Incorporating the constraints (2-16) and (2-17) produces the set of feasible contracts indicated in figures 2-5 and 2-6 (the shaded cone). Depending on the level of accuracy that the investor desires to implement, two possibilities arise. In the first case, the

---

[8]To find the corresponding solution for the case $\pi_L \geq \pi_H$, one simply interchange the subscript L and H.

tangency point T lies inside of the shaded cone of feasible contracts. In this case, the need to motivate the advisor is not in conflict with the need to elicit truthful revelation. The solution is then the same as the reduced problem. The advisor is paid a fixed amount whenever he makes a correct prediction. In the second case, the tangency point lies outside of the shaded cone. The optimal contract in the reduced problem is no longer incentive compatible. The optimal solution corresponds to the point at which the investor's iso-cost curve is just touching the shaded area. The two incentive concerns interact with one another and the constraint for truthful revelation becomes binding. As an intermediate case, the tangency point lies at the corner of the shaded area. The corresponding accuracy is formally stated in the following definition.

Definition 2.1: Let $\alpha=\lim_{\theta\to 1}\{\theta-\frac{C(\theta)}{C'(\theta)}\}$, we define

θ^0_H to be the solution to $\theta-\frac{C(\theta)}{C'(\theta)}=\pi_H$, if $\alpha>\pi_H>1/2$, $\theta^0_H=1/2$, if $\pi_H=1/2$, and

$\theta^0_H=+1$ if $\pi_H\geq\alpha$[9].

For an accuracy level $\theta<\theta^0_H$, the tangency point lies outside of the shaded area. For $\theta\geq\theta^0_H$, the tangency point lies inside of the shaded area. We summarize the optimal reward scheme in the following.

Lemma 2.4: Let $a_2(\theta)=(1-\theta)C'(\theta)+C(\theta)$ and $a_1(\theta)=\theta C'(\theta)-C(\theta)$, the optimal

---

[9]The existence of a finite α follows from that $\theta-\frac{C(\theta)}{C(\theta)}$ is both monotonic and bounded above. Since $C'(\theta=1/2)=0$, $\theta-\frac{C(\theta)}{C(\theta)}$ is defined at $\theta=1/2$ by continuous extension.

payment scheme implementing a given accuracy level θ is

i)if $\theta \geq \theta^0_H$ $w(x_H,r_H)=w(x_L,r_L)=h(C'(\theta))$;

ii)if $\theta<\theta^0_H$ $w(x_H,r_H)=h(\frac{a_1(\theta)}{\pi_H})<w(x_L,r_L)=h(\frac{a_2(\theta)}{\pi_L})$

Lemma 4.2 indicates that the critical accuracy θ^0_H separates the accuracy levels into two regions. The intuition behind this is the following. In exploiting the information advantage by exerting no effort and always reporting x_H (or x_L), the advisor loses the opportunity of profiting from correctly predicting the state at x_L (or x_H). Under a symmetric payment scheme (part i) of lemma 4.2), such loss in profit increases with the accuracy level induced. Above the critical accuracy level θ^0_H, this loss is larger than the advisor's cost saving from exerting no effort. Therefore, the symmetric payment scheme is sufficient to ensure that the advisor chooses the accuracy level induced and reports truthfully. Below θ^0_H, the symmetric payment scheme does not impose enough opportunity cost on the advisor. Consequently, it is insufficient to ensure that the advisor chooses the induced accuracy level and reports truthfully. The investor must reallocate the payment among the two instances when the advisor correctly predicates the states, so as to impose sufficient opportunity cost on the advisor. Given $\pi_H \geq \pi_L$, θ_H is more likely to occur based on the prior. Under a symmetric payment scheme, the advisor will always report x_H if he chooses to expend no effort. To prevent this, the investor decreases the payment when the advisor correctly predicts the favorable state and increases the payment when the advisor correctly predicts the unfavorable state. The new payment scheme thus increases the advisor's opportunity cost for not expending effort. In this case, the need to elicit truthful revelation interacts with the need to motivate the advisor to expend effort.

In the second step, the investor chooses the accuracy level to maximize his profit by employing the payment scheme specified in lemma 4.1 and 4.2. The following proposition summarizes the solution to the full problem. In the proposition, θ_{SB} denotes the second best accuracy level.

Proposition 2.2: At the solution to [I-P], $w(x_H,r_L)=w(x_L,r_H)=0$ and

i)if $\theta_{RP}\geq\theta^0_H$, then $\theta_{SB}=\theta_{RP}$ and $w(x_H,r_H)=w(x_L,r_L)=h(C'(\theta_{SB}))$;

ii)if $\theta^0_H>\theta_{RP}$ then the optimal compensation scheme is

$$w(x_H,r_H)=h(\frac{a_1(\theta_{SB})}{\pi_H})<w(x_L,r_L)=h(\frac{a_2(\theta_{SB})}{\pi_L})$$

and $\theta_{SB}<\theta^0_H$ solves the problem

$$Max_{\theta\in[1/2,\theta^0_H]}\{\theta\beta-\theta[\pi_H h(\frac{a_1(\theta)}{\pi_H})+\pi_L h(\frac{a_2(\theta)}{\pi_L})]\ .$$

Part i) of proposition 2.2 follows directly from lemma 2.4. If the optimal accuracy level induced in the reduced problem is higher than θ^0_H, the optimal payment scheme in the reduced problem is incentive compatible in the full problem. Clearly, it is also the optimal solution to the full problem.

Part ii) of proposition 2.2 indicates that, when $\theta^0_H>\theta_{RP}$, the optimal asymmetric payment scheme implementing θ_{SB}, in the region $(1/2,\theta^0_H)$, dominates all the symmetric payment schemes implementing accuracy levels higher than θ^0_H. The second best optimal payment scheme assumes the form indicated in part (ii) of lemma 4.2. In this case, the solution to the full problem is different from that in the reduced problem. The advisor is paid more when he correctly predicts the state less likely to occur.

Finally, we briefly discuss the problem of the uniqueness of the second best solution.

The second best solution is unique if $\theta_{RP} \geq \theta^0_H$. However, when the optimal payment scheme is asymmetric, i.e when $\theta_{RP} < \theta^0_H$, the investor's objective function is generally not concave[10]. The following tie-break rule guarantees the uniqueness of the second best solution.

Assumption 2.1: If there are multiple second best solutions, we assume that the investor will implement the solution with the highest accuracy level(largest θ).

The implication of the assumption is as follows. If there are two distinct θs both solving [I-P$''$], the advisor is strictly better off under the higher θ that solves [I-P$''$]. This follows from the fact that the expected total payoff to the investor's profit is the same under the two θs, and the fact that the investor's profit from investment is strictly increasing in θ. Therefore, assumption 2.1 amounts to assuming that when the investor is indifferent between implementing two different θ_{SB} he will implement the one most preferred by the advisor. Under assumption 2.1, the second best solution is unique.

Proposition 2.3: There exists a largest θ which solves the optimization problem in proposition 2.2 and hence the second best solution is unique under assumption 2.1.

The relations between the optimal accuracy levels are in general ambiguous. This is due to the dependence of θ_{SB} on the advisor's risk premium function. However, under the condition $h'''(\theta) \geq 0$ [11], the second best accuracy level is strictly smaller than the first best accuracy level.

Proposition 2.4: Given $h'''(\theta) \geq 0$, $\theta_{RP} < \theta_{FB}$.

---

[10]If $\pi_H = 139/144$ and $\theta_{RP} < \theta^0_H$, then the second order derivative of the objective function in proposition 2.2 is positive at $\theta = 3/4$ when the inverse utility function is $h(x) = x^2/2$.

[11]The third derivative of h(.) is positive when the advisor has constant or increasing risk aversion. It also holds when the advisor's risk aversion decreases slowly.

The following proposition indicates, under more restrictive conditions, the second best accuracy is also lower than that of the reduced problem.

Proposition 2.5: Assuming $h'''(\theta)\geq 0$,

i)If $\theta_{RP}\geq\theta^0_H$, then $\theta_{SB}=\theta_{RP}$

ii)If $\theta_{RP}<\theta^0_H$ and $\theta_{SB}\leq\pi_H$, then $\theta_{SB}<\theta_{RP}$.

When $\theta_{SB}>\pi_H$ and $\theta_{RP}<\theta^0_H$, θ_{SB} is not necessarily smaller than θ_{RP} as illustrated by the following example.

Example: Suppose the cost of information collection is $C_1(\theta)=\alpha(\theta-1/2)^2/2$. The investor's inverse utility function is $h(x)=x^2/2$, and the prior is $\pi_H=3/4$. For this case, we find $\theta^0_H=1$. From lemma 2.3, $\theta_{RP}<\text{Max}\{\theta^0_H,\theta^0_L\}$. For $\theta_{RP}=.9$, we find $\beta/\alpha^2=.44$. But for $\theta_{SB}=.9$, we find $\beta/\alpha^2=.42666$. Since θ_{RP} is strictly increasing in β/α^2. Thus, when $\beta/\alpha^2=.42666$, we have $\theta_{RP}<.9=\theta_{SB}$.

Our next result compares the investor's profits P_{FB}, P_{RP} and P_{SB} for the first best, the reduced problem and the second best respectively.

Proposition 2.6: i)The investor's profits are ordered by $P_{FB}>P_{RP}\geq P_{SB}$. The second inequality holds strictly when $\theta_{RP}<\theta^0_H$.

ii)Let $P_a(\pi_H)$ be the investor's profit in the second best solution. For fixed β, $P_a(\pi_H)$ is continuously decreasing in π_H, for $\frac{1}{2}\leq\pi_H<1$.[12]

Similar to the discrete model, the investor's profit decreases as the signal becomes the advisor's private information. This is due to the loss in risk sharing when the investor

---

[12]In changing π_H, we maintain the assumption that the expected payoff of the risky asset is the same as that of the risk free asset so that the optimal investment decision is unchanged.

must reallocate the payment between the two instances when the advisor correctly predicts the state.

The intuition for part ii) of proposition 2.6 is the following. From lemma 2.2, the advisor optimally exerts no effort and always reports either x_H or x_L if he deviates from the strategy induced by the contract. As prior π_H increases, the advisor is more certain about the state occurring. Therefore, it becomes more profitable for the advisor to exert no effort and predict the state most likely to occur. The cost of motivating the advisor increases. Consequently, the investor's profit decreases.

Extension

In the analysis so far, we have assumed that the ex post realization of state is costlessly observable. Sometimes return from investment not undertaken may never be observable. Such instances often arise for firms with firm-specific investment opportunities. To optimally compensate the manager, the assumption above must be relaxed.

In the following, we show that the analysis in the previous section can be directly applied to this case. To this end, we assume that the return from investment not undertaken can not be observed ex post. To facilitate comparison with results in the previous section, we assume $\pi_H \geq \pi_L$. Given the optimal investment decision, the risky project is not undertaken if an unfavorable signal x_L is reported. The assumption that the realization of state is not observed if the project is not undertaken implies that the compensation contingent on the unfavorable report must be independent of the state, i.e $w(x_L,r_L)=w(x_L,r_H)$.

The investor's problem in the extended model is similar to [I-P] except that we need to add the constraint $w(x_L,r_L)=w(x_L,r_H)$. In employing proposition 2.1 and lemma 2.2 to

reduce [I-P] to [I-P$^{\theta}$], it is not necessary to assume $w(x_L,r_L)\neq w(x_L,r_H)$. It follows immediately that the optimization problem corresponding to the extended model is the following.

$$Max_{\theta,\lambda(x),w(\cdot,\cdot)} E_{x,r}[W_0R+\lambda(x)(r-R)-w(x,r)|\theta]$$

$$E_{x,r}[V_A(w(x,r))|\theta]-C(\theta)\geq E_{x'}\{Max_{x'}E_r[V_A(w(x',r))|\theta=\frac{1}{2},x]|\theta=\frac{1}{2}\} \qquad (2\text{-}18)$$

$$\theta\in arg\max_{\theta'} E_{x,r}[V_A(w(x,r))|\theta'] \qquad (2\text{-}19)$$

$$w(x,r)\geq 0 \qquad x\in\{x_L,x_H\} \qquad r\in\{r_L,r_H\} \qquad (2\text{-}20)$$

$$w(x_L,r_L)=w(x_L,r_H) \qquad (2\text{-}21)$$

Again, we consider the problem of implementing a given θ at minimum cost. The following lemma describes such minimum cost payment scheme.

Lemma 2.5: The minimum cost payment scheme implementing accuracy level θ is $w(x_H,r_L)=0$,

$w(x_H,r_H)=a_1(\theta)/\pi_H+a_2(\theta)/\pi_L$, $w(x_L,r_L)=w(x_L,r_H)=a_2(\theta)/\pi_L$, where $a_1(\theta)$ and $a_2(\theta)$ are given in lemma 4.2.

The proof of lemma 2.5 proceeds in the same way as that of lemma 2.4.2, and is therefore omitted. Similar to the previous analysis, lemma 2.5 indicates that the payment is strictly positive at the unfavorable report. This is necessary to induce truthful revelation when the unfavorable signal is observed. In contrast to the previous model, the payment for correctly predicting good state is strictly higher than that at unfavorable report. When the state r_L is never observable if the project is not undertaken, the investor can not discern whether the manager's prediction is correct or not when he predicts the unfavorable state.

In this case, the only way in which the investor can verify manager's prediction is when the manager predicts the favorable state. Thus, the only indication that manager has expended effort is the correct prediction of favorable state. Consequently, he is rewarded more in such instances.

Conclusion

Our analysis focuses on how an investor best motivates a privately informed advisor to expend effort. The need to motivate the advisor stipulates that he should be rewarded more richly for correctly predicting the state if expending effort is most effective in enhancing his ability to predict that state. The presence of the unobservability of signal generally interferes with the need to motivate the advisor. Our analysis reveals that when the investor induces the advisor to acquire quite accurate information, whether the signal is publicly or privately observed is inconsequential. In other cases where the investor requires a lower accuracy level, inducing the advisor to reveal his information strictly increases the investor's cost of contracting with the advisor. Further, we find, under fairly general conditions, that the effect of moral hazard and hidden information in the investor--advisor relationship is to reduce the amount of effort that the advisor is induced to supply. The investor's profits also decrease if the investor is unable to monitor effort or to observe the investor's signal.

In considering directions for future research, we suspect that the methodology developed here might fruitfully be applied to examining other agency relationships. For instance, examples in which a company seeks the advice of a marketing expert on developing new sales strategy or a resource company consults with a geologist concerning

oil or minerals exploration have elements in common with investor--advisor relationship that we analyzed here.

Further extensions of our analysis would involve relaxing or modifying some of the simplifying assumptions we have employed. First, our binary information structure might be extended to allow for multiple signals and multiple states of nature. We suspect, however, that our main qualitative results, as summarized in Proposition 2.2, will continue to hold in the more general setting. Second, our single-investor and single-advisor relationship might be modified to allow for multiple advisors who supply the investor with independent assessments of investment prospects. A promising approach to modeling this case appears in the work of Dewatripont and Tirole (1995) on the use of advocates in agency relationships. Finally, our single period investor--advisor relationship might be modified to extend to several periods. Interesting issues arise in this setting as the advisor may modify his behavior to maintain or enhance his reputation. Reputational concerns may reduce the advisor's tendency to shirk or to misrepresent the signal he observes. These and other related issues await further research.

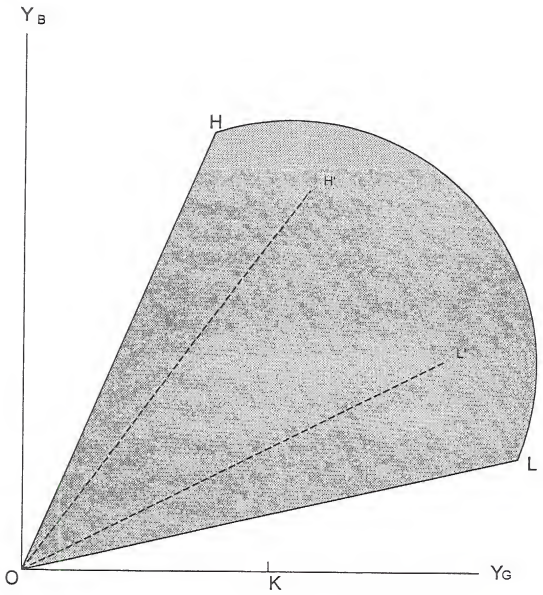


Figure 2-1

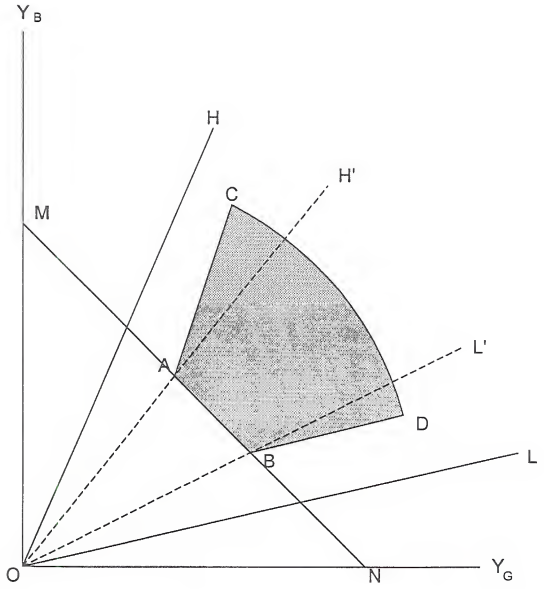


Figure 2-2

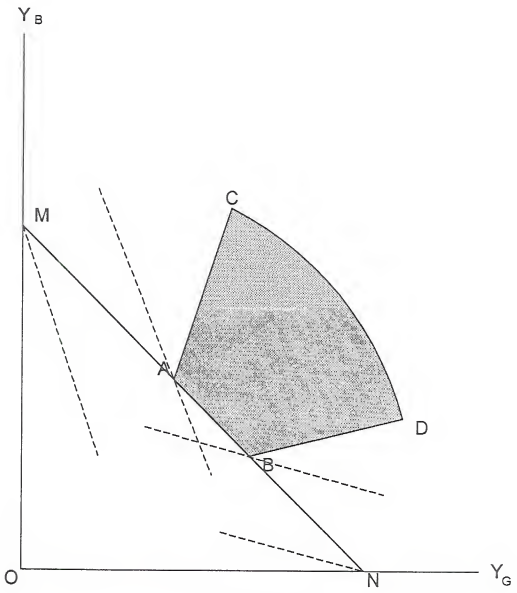


Figure 2-3

Slope of MN $-(p_g - q_g)/(p_b - q_b)$

Slope of iso-cost line (dotted line) $-p_g/p_b$

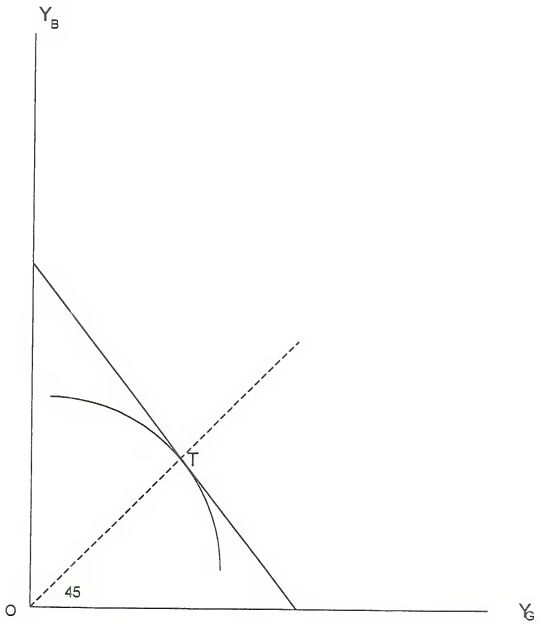


Figure 2-4

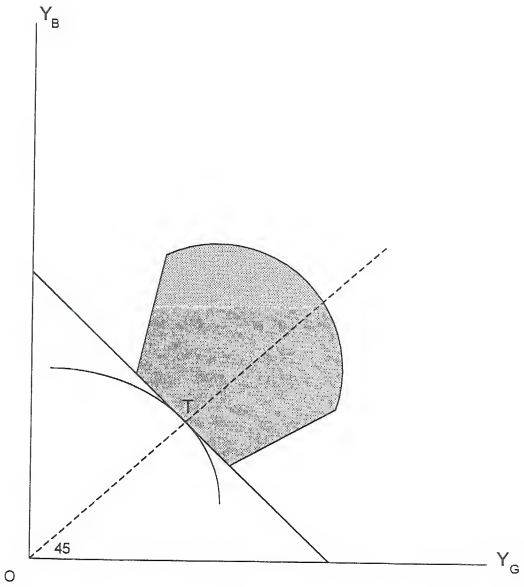


Figure 2-5

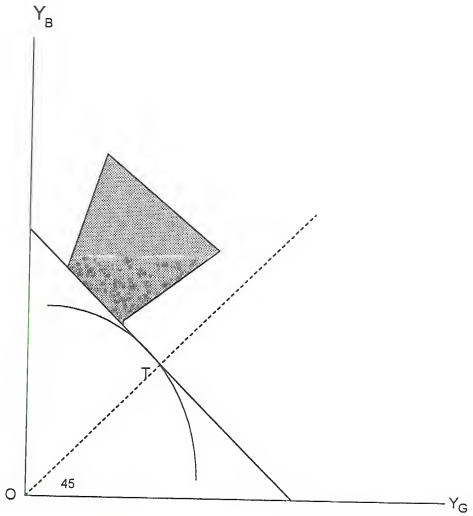


Figure 2-6

CHAPTER 3
SETTING LEGAL STANDARDS FOR CREDIBLE COMPLIANCE

Introduction

For most parties the threat of being fined or punished provides incentives to take care not to harm others. For instance, motorists may obey traffic regulations, industrial firms may resist fouling the air, and manufacturers may produce safe toys all to avoid fines for violation of standards.

The chance that a party will be fined not only depends on his action, but also on the effort that law enforcers exert to insure compliance. Recent experience reveals that it is difficult for public officials to control the behavior of enforcement agencies.[1] This suggests that law enforcers need to be motivated to detect violators, perhaps by rewarding them according to their success in discovering violations.[2]

In such a setting the equilibrium interaction between potential offenders and law enforcers will determine how regulations are observed and enforced. The amount of effort

---

[1]Most recently displeasure with the performance of the Internal Revenue's Service prompted Congress to cut the agency's compliance budget. Previous to this Congress had similarly intervened in the affairs of the FTC and the EPA to correct what it perceived as inappropriate enforcement of government policy.

[2]This approach differs significantly from most of the formal literature on law enforcement and monitoring, as exemplified by Baron and Besanko (1984), Border and Sobel (1987) and Mookherjee and P'ng (1992, 1994). These analyses assume that law enforcers can commit to a monitoring strategy independent of whether the strategy uncovers violators in equilibrium. A notable exception is Graetz et al (1986) who assume that enforcers are motivated by the fines they collect from prosecuting violators.

enforcers exert will depend on the perceived likelihood that parties have violated standards, and the likelihood of violation will depend on how vigorously the law is enforced. In turn the behavior of offenders and enforcers will be shaped by the standards determining if a party has violated the law. Examples of standards include a maximum number of product failures a manufacturer can experience before violating a safety code, or a minimum concentration of effluents found in a water sample that cause a waste discharger to violate emission regulations.

Beginning with Becker (1968) most analyses of the economics of enforcement have taken legal standards as given, and focused on the setting of fines as the primary tool of enforcement. In practice, though, the ability of enforcers to vary statutory fines is restricted by political, moral and legal constraints. In contrast, agencies may have some discretion in setting standards for determining when a party's actions are harmful. The primary goal of this chapter is to characterize how the setting of legal standards affects the behavior of complying parties, law enforcers, and the net social surplus generated by the regulation. Another goal of the chapter is to examine the extent to which setting standards and fines are substitute instruments for law enforcement.

Under optimal circumstances, where law enforcers can costlessly detect violations, offending parties should be induced to select care so that the marginal cost of care equals the social marginal benefit. However, we find that when enforcers must be incented to monitor compliance, it is desirable to induce care levels that either exceed or fall short of the surplus maximizing level.

The intuition for this finding is that some distortions in care are required to reduce the

cost of law enforcement. Suppose standards are initially set so that the marginal costs and benefits from taking care are equated. Then a slight variation in standards will not appreciably affect net benefits,[3] but it will cause a nontrivial adjustment in the enforcer's costs and effort. In some instances a slight loosening of standards will decrease enforcement costs. This will arise whenever looser standards causes enforcers to reduce their effort because the marginal returns from monitoring decrease as the probability of noncompliance decreases. We refer to this as the *complements* case because monitoring effort and standards are complementary inputs in determining the probability of a violation. In this instance, it will be desirable to loosen standards and induce less care in order to reduce the costs of enforcement.

For other applications monitoring effort may fall as the probability of noncompliance increases. This will arise if the returns from monitoring compliance in order to prove a violation will diminish as the degree of noncomplying behavior increases.[4] For this case, referred to as the *substitutes* case it will be desirable to set tighter standards and induce greater care in order to reduce the enforcer's expenditure on effort.

This is the central result of the chapter which is formally derived in Section 3. In Section 4 we consider the possibility that the costs of monitoring effort vary by the enforcer's ability to observe and process information. These costs are known privately by the enforcer. We show that the presence of asymmetric information reinforces our main finding that

---

[3]To a first order, a small change in standards has no effect on net benefits since marginal benefits and marginal costs of care are the same.

[4]For instance, it may not be necessary to expend much effort by employing sophisticated measuring devices to detect excessive discharge of effluents when polluters are in obvious violation of the law.

violation standards are distorted to reduce enforcement costs.

In section 5 we examine the possibility that parties differ in the costs they incur in taking care. We show how our main finding generalizes to this case, and demonstrate the optimality of allowing the highest cost parties to pay a fixed fee which absolves them from prosecution for a violation. Further, we demonstrate that corrupt enforcers can collude with potential offenders to similarly offer high cost parties protection from the law in exchange for a bribe.

In section 6 we examine the relationship between fines and standards. We find that , in contrast to Becker (1968), it is not necessarily desirable to impose the largest fine. Increases in fines may increase costly enforcement effort.

The chapter is concluded in section 7 with a summary of results and suggestions for further research. The elements of our model are introduced in the next section and all formal results are derived in the appendix. We relegate the discussion of related findings in the literature to those sections of the chapter where the results for comparison with the literature are presented.

Elements of the Basic Model

There exists a party who can exert some care denoted by $q > 0$ to avoid harming other individuals. For instance q, may be the discretion a motorist exercises to avoid an accident; q may be the control of emissions by a waste discharger, or q may be product quality a manufacturer supplies to avoid breakdowns. The party incurs a monetary cost or disutility of supplying q, denoted by $C(q)$ which is increasing and strictly convex with $C'(0) = 0$.

Social benefits from q are given by Bq, where $B > 0$, is the constant marginal benefit.[5]

The government sets a standard, denoted by s, as a criterion for determining if a party has exercised proper care. Depending on the application, s may be a speed limit which motorists must obey, or a maximum allowable concentration of pollutants in a discharger's water or air sample. To avoid the daunting task of explicitly modeling the bureaucratic and legal process by which violators are prosecuted we adopt a simpler reduced form description of the enforcement process. We assume that given s and q there is a probability that the party will be successfully cited for violating the standard denoted by $\tilde{P}(q,s,e) \in [0,1]$, where e is the effort the law enforcer supplies to monitor the party. We assume that this probability is decreasing as the party supplies more care at a decreasing rate with $\tilde{P}_q < 0$, and $\tilde{P}_{qq} > 0$ whenever e> 0. A tightening of standards increases the citation probability, $\tilde{P}_s > 0$ for $e >$ 0. Further $\tilde{P}$ is increasing in the enforcer's effort, at a decreasing rate so that $\tilde{P}_e > 0$, $\tilde{P}_{ee} < 0$. This implies that the burden of proof falls on the enforcer to demonstrate that a violation has occurred. Finally, we assume that the sign $(\tilde{P}_{es})$ = sign $(-\tilde{P}_{eq})$ which means that an increase in standards or a decrease in care both have the same qualitative effect on the enforcer's marginal returns from effort, $\tilde{P}_e$.[6]

As mentioned in the introduction, we distinguish between two cases describing how

---

[5]This specification of care benefits is made for simplicity and is not essential for the foregoing analysis.

[6]A simple specification that satisfies our assumptions is $\tilde{P}(q,s,e) = p(\delta,e)$ *where* $\delta = s - q$ measures the gap between the standard and the care provided, and $p_\delta, p_{\delta\delta} > 0$. In the context of pollution standards, δ might measure the difference between acceptable and actual effluent concentration in a water or air sample for example.

$d/ds\ (\tilde{P}_e) < 0$, an increase in standards affects the incentives for enforcers to monitor. In the *complements* case $d/ds\ (\tilde{P}_e) > 0$, and an increase in standards increases the marginal returns to monitoring. This might arise, for instance, if a party is cited whenever he is simultaneously violating the law and he is being monitored by the enforcer. In that case a tightening of standards will increase the probability that the party is in fact violating the law, which will therefore increase the enforcer's returns from monitoring. In the *substitutes* case, and a tightening of standards reduces the marginal returns to monitoring. This situation arises ,for example, if the enforcer knows whether a party has violated the law, but he must expend effort to prove the violation has occurred. When standards are tightened violations of the law are easier to demonstrate. Consequently, the enforcer's expenditure of effort required to prove a violation is reduced. [7]

---

[7] An example of a monitoring technology satisfying all the assumptions we have posited for the substitutes case is

$$\tilde{P}(q,s,e) = \begin{cases} \int_0^{s-\mu(q)} f(\lambda \mid e)d\lambda & s \geq \mu(q) \\ 0 & s < \mu(q) \end{cases}$$

where

$$f(\lambda \mid e) = B(e)e^{-B(e)\lambda} \quad ;\lambda \geq 0$$
$$B(e) = e/(1+e)$$
$$\mu(q) = \ln(1+q) \quad ;q > 0$$

In this example, a agent exercises care q to produce a product with quality $\mu\ (q)$. The enforcer observes a signal of quality, σ, given by

$$\sigma = \mu(q) + \lambda$$

Exerting greater effort allows the enforcer to observe quality with greater precision

If cited the party pays a fine, $F > 0$ for his offense. Consequently, the expected penalty for a violation is given by $P(q,s,e) \equiv F\tilde{P}(q,s,e)$. Throughout most of our analysis we assume that F is fixed, thus allowing us to focus on the setting of standards as the primary tool for shaping compliance and enforcement behavior.[8] Later in section 6 we examine the implications of varying the level of the fines, as well as the extent to which fines and standards are substitute instruments for law enforcement.

---

as reflected in the specification for $f(\lambda \mid e)$. One can easily verify that this specification satisfies our assumptions for the substitutes case

A slight variation on the first example allows us to produce another monitoring technology which satisfies all of our assumptions for the complements case. Here we assume that

$$\sigma = \mu(q) + \{1 - \exp[-(\tilde{\lambda} + g(e)/B(e))]\}$$

where

$$g(e) = -2\ln\left(\frac{e}{e+1}\right)$$

Then for $s \in (\mu(q) + 1 - e^{g(e)/B(e)},\ \mu(q) + 1)$

$$\tilde{P}(q,s,e) = \int_0^{\ln[1-(s-\mu(q))]-g(e)/B(e)} f(\lambda \mid e)\, d\lambda$$

which satisfies the assumptions required for the complements case.

[8]This treatment of fines differs from the economics of crime literature, as exemplified by Becker (1968), Stigler (1970), Polinsky and Shavell (1979), Malik (1990), Andreoni (1991) and Mookerherjee and P'ng (1992, 1994), which typically treats variations in fines as a primary enforcement tool. In reality the level of fines is set by the legislative branch, and the ability to adjust statutory penalties is restricted as noted by Graetz et al (1986). Harrington (1988) points out that the fines for violation of environmental standards are constrained to be quite small.

Enforcement of the standard is delegated to a single agency, who supplies effort to monitor potential offenders.[9] There is a cost borne by the agency personnel of supplying effort given by the function, *D(e)*, which is strictly increasing and convex in effort with $D'(0) = 0$. We make the realistic assumption that it is not possible for public officials to commit the agency to an enforcement policy or to know how diligently the agency enforces standards. Any agency model is likely to be deficient in describing some aspects of bureaucratic behavior, nonetheless we require some paradigm to proceed. We therefore assume that the agency selects an enforcement strategy to maximize the expected sum of fines collected net of the costs of enforcement effort.[10,11]

The interaction between the party and the enforcer is modeled as a game. The party chooses care *q(e;s)*, given the enforcer's effort and the standard where $q(e;s) = \underset{q}{\text{argmax}} \{U(q,e,s) = -P(q,s,e) - C(q)\}$. The enforcer chooses effort *e(q;s)* given the party's care decision and the standard, where

$$e(q;s) = \underset{\epsilon}{\text{argmax}} \{\Pi(q,e,s) = P(q,e,s) - D(e) + T\}.$$

---

[9]We are assuming that economies of scale in collecting and processing information dictate that enforcement be centralized.

[10]This approach is also employed by Graetz et al (1986) in their analysis of tax compliance. Our results do not change significantly if we assume more generally that the agency is rewarded based on some increasing function of the fines collected. For instance, promotion of agency personnel may be conditioned on their success at prosecuting violators.

[11]Alternatively, we might imagine that enforcement is undertaken by a private firm selected by the government. The relative advantages of employing private versus public law enforcement are discussed in Becker and Stigler (1974), Landes and Posner (1975) and Polinsky (1980).

T is a government transfer paid to the agency to insure it breaks even.[12] A Nash equilibrium to this game consists of a decision pair *{q(s), e(s)}* such that *q(s) = q(e(s);s)* and *e(s) = e(q(s);s)*. Below we demonstrate that such an equilibrium exists and that it is unique given *s*.

We assume that the government's objective function , $V = (Bq - T) + U + \lambda \Pi$, is the societal benefit of care net of government subsidies to the enforcer *(Bq-T)*, plus the utility of the party, U, plus the enforcer's profit, discounted by $\lambda < 1$. The discounting of enforcer profits derives from the fact that the government's primary constituency is the public at large, including the care providing parties.[13] In this case the government limits the agency's profit to zero. Rewriting V, the government's problem [G-P] becomes

$$\max V(s) = \max B(q(s)) - C(q(s)) - D(e(s)) \qquad \text{[G-P]}$$

The government selects a standard s to maximize the net benefit of inducing a given level of care, including the costs of enforcement given the Nash equilibrium behavior of the party and the enforcer.

Analysis of the Simple Case

For a given standard, s, the corresponding Nash equilibrium care level and enforcement effort are characterized by

$$-P_q(q,e,s) - C'(q) = 0 \qquad (3.1)$$

$$P_e(q,e,s) - D'(e) = 0 \qquad (3.2)$$

---

[12]Alternatively ,T is a tax which allows the government to collect excess revenues. when the agency generates positive profits.

[13]In the symmetric information case of section 3 the government sets T = P-D, so that $\Pi = 0$, and the government's objective function simplifies to become Bq-C -D

Given e, and s, the party selects care to equate the marginal reduction in expected fines to the marginal cost of care. The enforcer optimally responds to q and s by selecting effort to equate the increase in expected fines to the marginal cost of effort. Given our assumptions we have:

Proposition 3.1: A unique Nash equilibrium exists satisfying (3.1),(3.2)

The reaction functions for the party and the enforcer and the resulting Nash equilibrium for the case of complements and substitutes are displayed respectively in Figures 3-1 and 3-2. When the standard and enforcement effort are complements, an increase in care decreases the probability of noncompliance which causes the enforcer to allocate less effort as indicated by the negatively sloped reaction function *e(q:s)* in Figure 3-1. A decrease in enforcement effort induces less care as reflected by the positive slope of the *q(e:s)* reaction function. By contrast in the substitutes case, Figure 3-2 reveals that an increase in care induces greater effort from the enforcer, whereas greater enforcement effort causes the party to be less careful.[14]

The Nash equilibrium characterized by (3.1) and (3.2) corresponds to a given standard, s. To investigate how the equilibrium behavior of the party and enforcer vary with different standard levels we introduce the following assumption

Assumption 3.1: $$(dq/ds)\big|_{de=0} > (dq(e(s),s)/ds)\big|_{de(s)=0}$$

Assumption 3.1 provides sufficient conditions for determining how enforcement effort varies

[14]This result arises because the marginal reduction in expected fines from increasing care is decreased when enforcement effort is increased in the substitutes case.

with the tightness of the standards. To interpret this condition, note that $(dq(e(s),s)/ds)_{|de(s)=0}$ measures the response of care to an increase in standards required for the enforcer to maintain a constant level of effort. The expression $(dq/ds)_{|de=0}$ reflects the actual change in care for an increase in standards undertaken by the party assuming enforcement effort in unchanged. Assumption 1 requires that the actual change in care undertaken by the party is insufficient to maintain the enforcement effort at a constant level. This simply implies that a change in standards will induce a nonzero response from the enforcement agency. Assumption 3.1 is satisfied for the example where $P(q,e,s) = p(s-q,e)$[15].

The effect of tightening the standard on equilibrium care and enforcement is characterized by:

Proposition 3.2: A tightening of standards always leads to greater care. Given Assumption 3.1, tighter standards lead to more enforcement effort in the complements case, and it leads to less effort in the substitutes case.

According to Proposition 3.2, the party always increases care as standards tighten to partially reduce the probability of being cited. Despite this increase in care, the opportunity for the enforcer to find a violation increases with a tightening of standards. This leads to an increase in effort when standards and effort are complements as the enforcer's marginal return from effort increases. In contrast, when effort and standards are substitutes the enforcer reduces effort since there is less need for monitoring to convict the party.

---

[15]When

$$P(q,e,s)=P(s-q,e)=P(\delta,e),\ \textit{then}\ (dq/ds)_{|de=0} = 1 > P_{\delta\delta}/(P_{\delta\delta} + C^{ee}) = (dq(e(s),s)/ds)_{|de(s)=0}$$

The government sets a standard to maximize the net benefits from care, including enforcement costs. If enforcement were costless, it would be optimal to set standards to induce care levels which equate the marginal benefit and marginal cost of care. This prescription for setting standards will not be optimal, however, when enforcement is costly. For suppose we begin with such a standard and assume that effort and standards are complements. A small reduction in standards. will decrease care, but, there will be virtually no effect on net benefits since the marginal benefits and marginal costs of care are approximately equal. However, a small reduction in standards will cause enforcement effort costs to decrease by a non negligible amount. Consequently a small reduction in standards below the level which would cause the marginal benefits and costs of care to be equated, will result in an increase in net surplus inclusive of compliance costs. A similar argument establishes that when standards and enforcement effort are substitutes, it is optimal to increase standards above the level which would induce the net benefit maximizing level of care. This is the intuition underlying the following proposition. In that proposition we refer to q^* as the care level which maximizes the net benefits from care (excluding enforcement costs) and $s(q^*)$ as the standard which induces q^* in equilibrium.

Proposition 3.3: Let $\tilde{s}$ be the solution to [GP]. In the complements case, $\tilde{s} < s(q*)$ and $B - C'(q(\tilde{s})) > 0$ as the optimal standard induces less than the net surplus maximizing level of care. In the substitutes case, $\tilde{s} > s(q*)$ and $B - C'(q(\tilde{s})) < 0$ as the optimal standard induces more than the net surplus maximizing level of care.

Proposition 3.3 shows how the enforcement monitoring technology influences the standards for due care, as well as the care level provided in equilibrium. When standards and effort are

complements, then standards must be relaxed to prevent enforcers from being overzealous in ensuring compliance. This could possibly explain why some safety and environmental standards appear to be too lax from the view point of the general public. Landes and Posner (1975) have similarly noted that it may be necessary to reduce violation fines to prevent over investment by private enforcers.

The results for the substitutes case are perhaps more surprising. One's intuition might suggest that when enforcement is costly this would add to the costs of inducing parties to take care thus making it optimal to induce lower care. However in the substitutes case, compliance costs are reduced by making it easier for enforcers to convict parties by tightening the standards, but tighter standards induce the parties to supply greater care.

Privately Informed Enforcer

In this section we extend our basic model to consider instances in which the enforcer's cost of effort is private knowledge. Such cases may arise when the cost of monitoring varies by the diligence required to apprehend offenders, by the nature of the offense, or by the characteristics of the parties. All of these attributes may be privately known by the enforcement agency. Hidden information may present difficulties for the government, if it operates under a fixed budget, and the agency claims its costs of enforcement are high. The government must insure the agency staff are adequately compensated to insure their participation, but it also must minimize the expenditures required to run the agency. We focus here on how care standards are optimally set under these circumstances.[16]

---

[16]To our knowledge the impact of privately informed enforcers on the design of optimal fines and standards has not been analyzed in the literature.

Suppose that the cost of effort is given by $D(e, \theta)$ where θ is a cost parameter known privately by the enforcer, with the properties that $D_{\theta}(e, \theta)$, $D_{e\theta}(e, \theta) > 0$ so that total cost and marginal cost of enforcement are increasing in θ.[17] The government is unaware of the realization of θ, but it knows that θ is distributed according to the density $f(\theta) >$ for $\theta \varepsilon [\underline{\theta}, \overline{\theta}]$.

We assume that the timing of the interaction between the government, the agency and the party is: first, the agency observes θ. Second, the government offers the agency a menu of contracts $\{T(\theta), s(\theta)\}$, where the dependence of the pair on θ, denotes that it is intended for the agency of type θ.[18] T is a reimbursement paid by the government to the agency to help cover its enforcement expenses. Third, the agency selects a preferred contract. The contract choice is public knowledge and the parties update their beliefs about the type of the enforcer based on the agency's contract choice. Fourth, simultaneously the parties choose their level of care, and the agency selects enforcement effort. Finally the agency collects fines from those parties found to be in violation of the standard.

Let $\Pi(\theta'|\theta)$ denote the agency's expected profit who selects the contract $\{T(\theta'), s(\theta')\}$ when their type is θ, where

$$\Pi(\theta'|\theta) = P(q(s(\theta'),\theta'), e(s(\theta'),\theta), s(\theta')) - D(e(s(\theta'),\theta),\theta) - T((\theta'),$$

$q(s(\theta'),\theta')$ is the equilibrium care level for the standard $s(\theta')$ given that the enforcer has chosen the contract intended for type θ'. The enforcer's contract choice affects the parties' beliefs about the enforcer which influences their choice of care. The equilibrium enforcement

---

[17]We continue to assume that D is increasing and strictly convex in e, and that $D_e(0,\theta) = 0$.

[18]That is, the menu is designed so that type θ will choose $\{T(\theta), s(\theta)\}$

effort $e(s(\theta'),\theta)$ depends on the standard, as well as on θ which is the enforcer's type.

The government's problem [GP-A] for this case is to choose $\{T(\theta), s(\theta)\}$ to

$$\max\ E_\theta\ \hat{V}(s(\theta),\theta)) \qquad \text{[GP-A]}$$

where E_θ is the expectation taken with respect to θ, and such that for all $\theta\varepsilon[\underline{\theta},\overline{\theta}]$: (i) the agency breaks even, $\Pi(\theta) \equiv \Pi(\theta|\theta) \geq 0$, (ii) the party picks the contract which is intended for it, $\Pi(\theta|\theta) \geq \Pi(\theta'|\theta)$.

In what follows, we focus on the separating equilibria solution to [GP-A] in which each type θ is induced to select a separate contract.[19] As a convenient benchmark for this solution to [GP-A] consider the complete information case, analyzed in section 3, where the government and the party know the agency's cost parameter, θ, at the time of contracting. Let $s^*(\theta)$ be the standard which induces the party to choose the net benefit maximizing care, q^*, in equilibrium. Refer to $\tilde{s}(\theta)$ as the optimal standard given the agency is known to be of type, θ. We then have :

Proposition 3.4: In the separating solution to [GP-A] the optimal standard, $\hat{s}(\theta)$ satisfies (i) $\hat{s}(\theta) \leq \tilde{s}(\theta) \leq s^*(\theta)$ for the complements case, and (ii) $\hat{s}(\theta) \geq \tilde{s}(\theta) \geq s^*(\theta)$ for the substitutes case (with strict inequality for $\theta > \underline{\theta}$ in both cases).

The presence of a privately informed agency causes a greater distortion in standards away

[19]Another possible policy for the government is to offer pooling or semi-pooling contracts in which several different types of enforcers are induced to accept the same contract. In this case, the enforcer's choice of a contract would not necessarily reveal his type. Such a policy might be beneficial if it were less costly to enforce standards when the enforcer's type was not known by the care providers. Deriving conditions under which pooling or separating contracts are preferred seems quite difficult, and therefore determining the optimal form of contract remains an open question. Although we focus on separating contracts in our discussion, we demonstrate in the appendix that Proposition 3.4 also holds for the case of pooling contracts

from $s^*(\theta)$, the level which induces the net benefit maximizing care. This arises because the agency will try to overstate its costs to obtain a more favorable contract from the government. In the case of complements the government reacts by reducing compliance standards which decreases the enforcer's effort. This renders it less attractive for a low cost enforcer to claim to be high cost, by reducing the number of effort units over which he can exercise his cost advantage. As a result of the reduction in standards the party provides less care as $q(\hat{s}(\theta)) < q(\tilde{s}(\theta)) < q^*$.

When effort and care are substitutes the government increases the standards, thus reducing the incentives for the enforcer to monitor. Again this makes it less attractive for a low cost enforcer to pretend to be high cost, because it reduces the number of effort units over which he may exercise his cost advantage. This tightening of standards induces the party to increase its care as $q(\hat{s}(\theta)) > q(\tilde{s}(\theta)) > q^*$.

Heterogenous Parties

In this section we examine desired alterations in optimal standards when there is a heterogenous population of parties varying according to their cost of taking care. Variations in cost arise because the parties have access to different methods to reduce the harmful effects of their behavior[20]. Further we assume that the parties are privately informed about their cost of taking care. As in the previous cases we've studied, the government sets a uniform standard which parties must adhere to. However, with a heterogenous population, the

[20]For instance, firms may differ according to the costs they incur to reduce pollution.

government may grant higher cost parties immunity from the standard, if they pay a fix fee.[21] This arrangement saves high cost parties the expense of meeting standards, while reducing the enforcer's monitoring costs.[22]

We model the heterogenous party population by assuming that an individual's cost of care is given by $C(q,\mu)$, where μ is a privately observed cost parameter. Total and marginal costs are increasing in μ, *with*, $C_{\mu}, C_{\mu q} > 0$, for $q > 0$.[23] The density of parties of type μ in the population, which is normalized to one is given by $g(\mu) > 0$ for $\mu \varepsilon [\underline{\mu}, \overline{\mu}]$. We assume the government offers parties the choice of either paying a fixed assessment, A to the enforcer, which exempts them from being cited, or the choice of trying to meet the standard*s*. Let $q(s,\mu) = \text{argmax}(-P(e(s),q(s),s) - C(q,\mu))$, be party type μ's optimal care to avoid being fined. Given A, and *q*, type μ's response is to pay A and avoid providing care if $(-P(q(s,\mu)e(s),s) - C(q(s,\mu),\mu)) \leq -A$, otherwise the party provides care $q(s,\mu)$. For a given A, some subset of the highest cost individuals $\mu \varepsilon (\hat{\mu}, \overline{\mu}]$ for $\hat{\mu} < \overline{\mu}$ will elect to pay the assessment, A. The cutoff type, $\hat{\mu}$ will just be indifferent between investing in care and paying the assessment to avoid being cited.

The government's problem, for the case of heterogenous parties, [GP-P] is to choose

---

[21]Alternatively, parties may self report their violations to the agency, where upon they are assessed a fixed fee. as in Kaplow and Shavell (1994).

[22]In theory if the set of potential offenders was known by the government, a menu of different standards and fines could be offered to separate out offenders by their cost of taking care. This approach is employed by Mookherjee and P'ng(1994) in their analysis of marginal deterrence of crime. Such fine tuning of standards is impractical however when the identity of the offenders is unknown at the time standards are determined.

[23]We continue to assume that C is increasing and strictly convex in q with $C_q(0,\mu) = 0$

the assessment A to

$$\max E_{\mu < \hat{\mu}} \{Bq(s,\mu) - C(q(\mu,s),\mu)\} - D(F(\hat{\mu})e(s)) \qquad \text{[GP-P]}$$

The maximand in [GP-P] represents the expected net benefit of care minus the enforcement costs taken over the population of parties investing in positive care levels. Those parties $\mu > \hat{\mu}$ who exempt themselves, contribute zero net benefits and impose zero enforcement costs on society. The solution to [GP-P] is characterized in the following proposition. In that proposition we refer to $\tilde{s}$ as the optimal standard, and s^* as the standard that maximizes $E_{\mu<\hat{\mu}}\{Bq(s,\mu) - C(q(s,\mu),\mu)\}$

Proposition 3.5: In the solution to [GP-P] (i) no parties are exempted from standards when B is sufficiently large, (ii) when exemption occurs A < F, and $\hat{\mu}$ satisfies $v(\hat{\mu}) = Bq(\tilde{s},\hat{\mu}) - C(q(\tilde{s},\hat{\mu}),\hat{\mu}) - D'(F(\hat{\mu})e(\tilde{s}))e(\tilde{s}) = 0$ (iii) $\tilde{s} < s^*$ for the case of complements, and (iv) $\tilde{s} > s^*$ for the case of substitutes.

Part (i) of Proposition 3.5 indicates that parties are exempted only if the benefits from taking care are sufficiently small, otherwise even high cost care providers are induced to provide care. Part (ii) indicates when exemption arises that higher cost parties opt to pay the assessment rather than risk paying a higher fine if they are cited. The assessment is set at a level so that only those parties with a negative care contribution to social welfare , net of marginal enforcement costs, $v(\mu)$, seek exemption. Parts (iii) and (iv) verify that the same distortion in standards arises when parties are heterogenous as when they are homogenous.

When exemptions are possible, dishonest enforcers may also take bribes from parties not wanting to provide care. To analyze this possibility, suppose for now that government sanctioned exemptions are not offered, perhaps because the benefits from care are too large.

Imagine that the enforcer offers any party an exemption from being monitored if the party pays the enforcer a bribe equal to Y. Assume also that such illegal activity goes unnoticed by the government, and that agreements between parties and the enforcer are kept. [24] Given the standard, s, the enforcer's problem , [EP] is to set the level of the bribe, Y and enforcement effort $e(s)$ to

$$\max E_{\mu < \mu'}\{P(q(s,\mu),e,s)\} - D(F(\mu')e) + (1-F(\mu'))Y + T \qquad [EU]$$

where all parties $\mu \varepsilon (\mu', \overline{\mu}]$ pay the bribe and type μ' is indifferent to paying the bribe and investing in care. The solution to the enforcer's problem is characterized in ,

Proposition 3.6: In the solution to [EP], (i) the enforcer always offers a bribe $Y < F$ which the higher cost parties $\mu \varepsilon (\mu', \overline{\mu}]$ pay. (ii) Y satisfies

$$(1 - F(\mu') - Y(d\mu'/dY)f(\mu')) = -\{P(q(s,\mu'),e(s),s) - D'(F(\mu')e(s))e(s)\}(d\mu'/dY)f(\mu')$$

According to Proposition 6 the enforcer always offers a bribe which some non neglible subset of the higher cost parties agree to pay for exempting themselves from being cited. The optimal bribe, characterized by the equality in (ii) sets the enforcer's marginal revenue from an increase in the bribe to the marginal increase in the collection of fines as more types invest in care in response to an increase in the bribe.

Propositions 3.5 and 3.6 suggest that if illegal bribes cannot be detected, high cost parties will always exempt themselves from fines by paying the enforcer a fee. In cases where the benefits from care are large, the fee will be a bribe paid to the enforcer, as assessments for exemptions will not be sanctioned by the government. In cases where the benefits from care

---

[24]One rationale for why corrupt agents may trust one another to honor agreements is that they may want to maintain a reputation for being reliable. See Tirole (1992) for one approach to modeling collusion between corrupt individuals.

are small, the fee may be a government sanctioned assessment, if A is less than Y. [25]

Setting Optimal Fines

To this point in our analysis we have assumed the level of fine for a violation, F, is fixed exogenously. Here we investigate whether increases in F are welfare improving. Becker (1968) first observed that larger fines deter parties from breaking the law and thus reduce enforcement effort required to insure compliance. As we demonstrate, this argument may fail to apply when the enforcer's effort supply depends on the probability that the party is in compliance.[26]

Suppose the fine, F, is increased. This will cause the government to adjust its optimal standard, s, and it will induce both the party and the enforcer to adjust their behavior. Let de/dF and dq/dF represent respectively the rate of change in equilibrium enforcement effort and care as F is increased. Then the increase in welfare from a change in F can be written as

$$dV/dF = (B-C_q)(dq/dF) - D_e(de/dF)$$

$$= \{(B-C_q)/D_e - (de/dF)/(dq/dF)\}D_e\,(dq/dF)$$

$$\left(\gtrless\right) 0 \;\; as \;\; (de/ds)/(dq/ds) \quad \left(\gtrless\right) \; (de/dF)\,/(dq/dF) \tag{3-3}$$

where the first line of (3-3) follows from the Envelope Theorem, the second line follows from

---

[25]We conjecture that A will be less than Y for B sufficiently small, although we have so far been unable to verify this.

[26]Several analyses have discovered reasons why maximal fines may be not be desired. Malik (1990) demonstrates that increasing fines may increase agent's avoidance behavior, thus leading to higher enforcement costs. Andreoni (1992) argues that juries are less apt to convict offenders when fines are more severe. thus reducing the deterrence power of maximal fines. Polinsky and Shavell (1979) argue that maximal fines are welfare decreasing in that some offenses should not be deterred if marginal benefits of the crime exceed the marginal costs. Stigler(1970) and Mookherjee and P'ng (1994) show that fines should be varied continuously in order to maintain marginal deterrence in enforcement.

the first by rearranging terms and the last line follows from the condition for setting optimal standards, *(dV/ds = 0)*[27]

A necessary and sufficient condition for ordering *(de/ds)/(dq/ds)* and *(de/dF)/(dq/dF)* and thus determining whether increasing fines is welfare enhancing is given in

Proposition 3.7: $(de/ds)/(dq/ds) \left(\gtreqless\right) (de/dF)/(dq/dF)$ *as* $d/ds \{-P_e/P_q\} \left(\gtreqless\right) 0$.

To interpret (3-3) note that under the optimal standard *(de/ds)/(dq/ds)* represents the rate at which enforcement effort and care may vary while keeping total surplus constant. In the complements case, too little care is allocated. An increase in F will induce the party to provide more care, but it will also cause the enforcer to expend more effort. If the rate at which extra effort expended for an increase in care is sufficiently small (less than *(de/ds)/(dq/ds)*) then increasing the fine will increase welfare. Otherwise increasing the fine will reduce welfare, if it will induce too much enforcement effort to be expended. A similar argument serves to confirm this intuition for the case of substitutes.

Proposition 3.7 provides necessary and sufficient conditions for an increase in the fine to be welfare decreasing. It's easy to verify that in the substitutes case where $P_{es} < 0$, that $d/ds\{-P_e/P_q\} < 0$. This implies that a small increase in the violation fine is welfare decreasing and it provides an interesting exception to Becker's argument for maximal fines. The intuition supporting this finding is that in the substitutes case, the level of care induced is excessive in order to limit enforcement effort. (see Proposition 3.3) An increase in the fine reduces welfare, by causing parties to further increase care which also induces enforcers to expend more effort.

---

[27]The optimal standard satisfies $dV/ds = 0$ or $B\text{-}C_q)/D_\varepsilon = (de/ds)/(dq/ds)$.

Conclusion

Our analysis offers one rationale for the divergence between the marginal benefits and the marginal costs from taking care which often arise in practice. Pollution and safety standards may either be set too loose or too stringent to discourage enforcers from exerting excess effort. Whether standards are set too low or too high depends on the available technology for identifying violators.

Our analysis also reveals the importance of setting standards, not only to influence compliance, but also to shape the behavior of enforcers. In circumstances where penalties are fixed, varying standards may be one of the few tools policy makers have to affect compliance and reduce enforcement expenses. In instances where fines can be varied as well, it may be counterproductive to set maximal fines which encourage overzealous law enforcement.

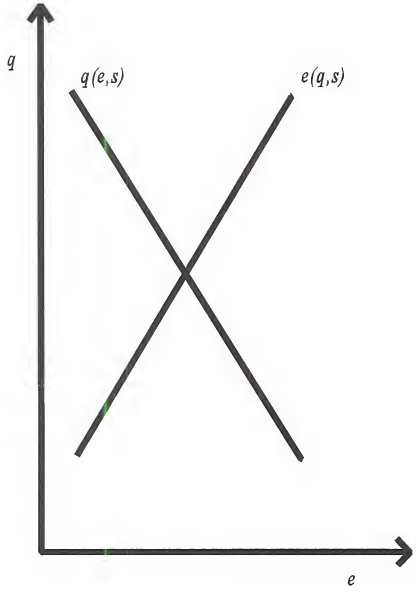


Figure 3-1
Complements Case: $P_{es}>0$

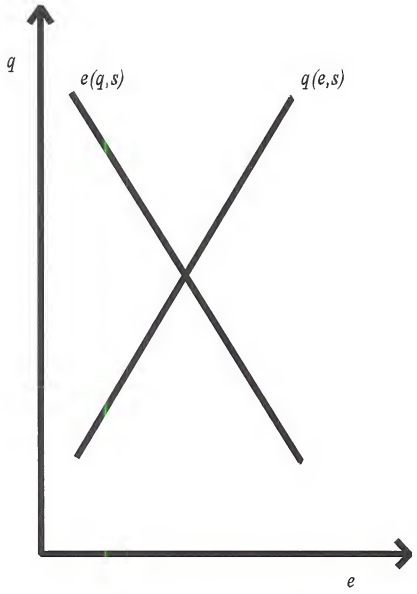


Figure 3-2
Complements Case: $P_{cs}<0$

CHAPTER 4
MONITORING AND THE OPTIMAL MIX OF PUBLIC AND PRIVATE DEBT CLAIMS

Introduction

Two of the most important functions of intermediated lending are that it facilitates delegated monitoring and provides flexibility. The theory of finance suggests that diverse groups of public lenders do not sufficiently monitor the firms to which they provide funds: It points to free rider problems and inefficient monitoring technologies as contributing reasons. In contrast, financial intermediaries are often assumed to have a unique advantage in performing monitoring. Intermediaries can act as delegated monitors by monitoring and controlling borrowers on behalf of other lenders, and in turn produce information[1]. Moreover, conflicts of interests and legal restrictions render negotiations with dispersed public lenders very costly, if not impossible. In contrast, concentrated lending by intermediaries enables negotiations to occur at ease, and provides flexibility allowing modifications of loan contracts as circumstances necessitate.

Despite the success in our understandings of intermediated lending, important questions remain unanswered (or only partially answered): 1) Given intermediaries' advantages in performing monitoring, what is their incentive to actually provide monitoring?

[1] See Berlin and Loeys (1988), Campbell and Kracaw (1980) and Diamond (1984). Empirical evidences based on loan announcement are documented in Billet, Flannery and Garfinkel (1995), James (1987) and Lummer and McConnell (1989). Fama (1985) provides another piece of evidence.

2) Casual observations of firms' capital structures reveal that they frequently borrow from both intermediaries and public debt market. Given the advantages of intermediated lending, why do firms demand public lending as well? 3) The notion that intermediaries act as delegated monitors assumes that public debtors' incentives over monitoring are in accordance with that of intermediaries'. Given that firms choose their debt structures, why do they desire to align the two incentives? This paper provides some answers to these questions.

To analyze these issues, we employ the following model: A wealth constrained manager must seek external financing to start a project, which lasts two periods and produces cash flows only on the final date. The distribution of the final cash flow depends on the interim states which are privately observed by the manager. The project provides positive NPV only in the favorable state. There are two sources of external financing, public lending and intermediated lending (private lending). Costly monitoring and costless interim renegotiation are feasible only with intermediated lending. Besides being the residual claimant of the cash flow from the project, the manager enjoys non-transferable private control rents, provided the project is continued at interim. The existence of control rents is the source of potential misalignment of preferences between the manager and the lenders over the interim continuation decision: In pursuit of the control rents, the manager may prefer to continue the project even if liquidation benefits the lenders. We assume that for lenders to break even, interim liquidation in the unfavorable state must occur with positive probability and it must strictly benefit the lenders.

To raise initial financing, the firm must assure lenders of an expected repayment equal to the amount of fund lent. The divergence of preference, between lenders and the manager,

over the interim continuation decision implies that monitoring may be required to ensure timely liquidation and initial financing. Our analysis is based on the observation that while monitoring can facilitate initial financing by mitigating agency problems, it introduces deadweight costs: 1)There is a cost for expending monitoring effort; 2)Liquidation destroys the manager's control rent. Thus, in designing the optimal debt structure, the manager has two goals: 1)To credibly payout the cash flows so that the required level of monitoring for initial financing is minimized; 2) To structure the private debt claim to induce the required level of monitoring. The main result of this analysis is that in general the manager can not achieve both of his two goals by relying entirely on private debt financing, and the optimal debt structure is a mix of both public and private debt. To derive this result, we proceed in several steps.

We first analyze the case in which the manager's private rent is sufficiently small. In this case, liquidation in the unfavorable state generates positive surplus, and therefore is efficient. We find that if the firm borrows long-term bank debt[2], which requires a repayment only after the cash flow from the project is realized, then in the unfavorable state the project is liquidated through renegotiation independent of whether the bank is informed or not. The need to borrow public debt arises because the bank's debt claim must be renegotiated to induce liquidation, so the division of surplus from liquidation can not be specified through ex ante contracting. If the bank is the only lender, the firm can extract most of the surplus from liquidation when it commands large bargaining power, and initial financing may become

[2]Henceforth, we will use the term "bank" to represent all types of institutions which can provide similar functions in our model. These institutions may include insurance companies, pension funds, etc.

infeasible. Unlike the bank lender, however, public debtors can free ride on the benefit of negotiations without making concessions of their claims in liquidation. By equating public debtors' claims in liquidation to the surplus generated, the firm can credibly pay out the surplus and enable initial financing without monitoring.

When the private rent is sufficiently large, the manager never desires to liquidate the project. Feasibility of initial financing requires involuntary liquidation in the unfavorable state. Therefore, the firm's initial borrowing must include bank debt requiring a repayment when the interim state is realized. Such a debt claim confers the interim control rights upon the bank, allowing it to force liquidation. Given the control rights, the bank can benefit from better information which enables it to timely liquidate the project in the unfavorable state. Thus, the bank's desire to maximize the value of the control rights motivates it to monitor.

When the manager raises initial financing only from a bank, we find that the optimal bank debt requires repayments both at interim and on the final date. Moreover, the size of the final repayment depends on the relative bargaining power between the manager and the bank lender. When the manager commands larger bargaining power, initial financing requires that the final repayment be sufficiently large. If the final repayment is small, the bank never forgives the interim repayment and the project can only be continued through renegotiation. When the manager commands large bargaining power, he can extract most of the surplus from continuation, and initial financing becomes infeasible. Increasing the promised final repayment increases the bank's expected payoff. This follows because if renegotiation breaks down, the bank can choose between its payoff in liquidation and its payoff when it forgives the interim repayment and allow the project to continue. Increasing the promised final

repayment raises the bank's reservation level, and therefore increases its expected payoff from continuation. When the bank commands large bargaining power, initial financing is always ensured. In this case, the manager desires to control the bank's benefit from being informed in order to reduce the costs associated with monitoring. We show that there exists a bank debt claim so that the bank's payoff, when it is informed of the favorable state, is equal to the promised interim repayment. By optimally setting this repayment, the firm can reduce the amount of bank monitoring.

In general, however, it is strictly suboptimal for the firm to raise initial financing only form a bank. To minimize the level of monitoring required by initial financing, the manager desires to maximize the bank's benefit per unit of monitoring effort. On the other hand, to induce the minimum level of monitoring, the manager must structure the bank's debt claim to control its benefit from monitoring. Thus, when the project is financed only by a bank, the manager's two goals are in conflict with each other.

In addition to bank debt, the firm can raise initial financing by also borrowing public debt. We consider the two cases in which the firm borrows either long-term or short-term public debt. In both cases, we find that by financing the project with a mix of public and bank debt, the firm can separate its two goals in designing the optimal debt structure: It can regulate the bank's incentive to monitor without interfering its desire to minimize the required level of monitoring for initial financing. To minimize the required level of monitoring, the manager desires to maximize the lenders' (the bank's and the public lenders') total expected payoff when the bank is uninformed. If the project is financed by a mix of public and private debt claims, then, for any fixed payoff scheme for the bank, the manager can maximize this

payoff by paying out cash flow form the project to the public debtors. Furthermore, by giving the public debtors a share of the proceeds from the liquidation, the manager can control the bank's benefit from monitoring, so that the desired level of monitoring is induced.

In comparing the optimal mix of bank debt and long-term public debt and that of bank debt and short-term public debt, we find that the former strictly dominates the latter. With short-term debt claims, the public lenders are repaid in full whenever the project is refinanced and allowed to continue by the bank. In this case, the public lenders can never strictly benefit from timely liquidation in the unfavorable state, which arises only when the bank is informed. In contrast, with long-term debt claim, the public lenders' payoff is state contingent and they can benefit from timely liquidation. Thus, long-term public debt claim allows the firm to align the public debtors' incentive over monitoring with that of the bank's, so that the debtors' total marginal benefit from monitoring is maximized, further reducing the required level of monitoring for initial financing. When the project is financed by the optimal mix, the bank acts as a delegated monitor.

There is an extensive literature on optimal debt structures. Our analysis is most closely related to that of Hart and Moore (1991). We share similar premises that managers design debt structures to credibly assure the lenders of their repayments. The main difference between the two is that Hart and Moore assume that there is no asymmetric information and renegotiation is frictionless. Therefore, there are no differences between public lending and bank lending in their analysis. Our analysis is also related to those by Rajan (1992), Park (1994), and Rajan and Winton (1995). Rajan (1992) analyzes the bank hold-up problem. He argues that firms can use public debt to mitigate the distortion in managers' incentives to

expend effort caused by banks' opportunism. Both Park (1994) and Rajan and Winton (1995) demonstrate that optimal enforcement of debt covenant can provide banks with incentive to monitor. Our analysis, on the other hand, is based entirely on the distribution of cash flows. Diamond (1991, 1993) investigates how firms having private information choose their debt structures. In our analysis, prior information is symmetric among agents.

The rest of the paper is organized as follows. In section I, we outline our model. In section II, we consider the case in which the manager's private rent is sufficiently small. In section II, we analyze the bank's incentive to monitor and derive the optimal bank debt. In section IV, we derive the optimal mix. Section V presents empirical evidence and section VI concludes.

Elements of The Model

There are three dates, t=0,1,2. There is an indivisible project which requires an initial investment of I_0 at t=0. The project returns a stochastic cash flow of r at t=2 distributed over the compact support [0,X]. The distribution of r is denoted as $F(r|\theta)$ and depends on the interim state $\theta \in \{\theta_H, \theta_L\}$. $F(r|\theta_H)$ strictly dominates $F(r|\theta_L)$ according to first-order stochastic dominance. If the project is terminated at t=1, the firm's assets can be liquidated at $L<I_0$. The terminal value of the assets at t=2 is zero. The parameters satisfy

Assumption 4.0: $$\int_0^X rdF(r|\theta_H) > I_0 > L > \int_0^X rdF(r|\theta_L). \tag{4-1}$$

Assumption 4.1: $$L \leq E_0[\int_0^X rdF(r|\theta)]. \tag{4-2}$$

Assumption 4.0 indicates that the project provides positive NPV in the favorable state while liquidation yields more cash flow in the unfavorable state. Assumption 4.1 says that, at t=1,

continuing the project is more profitable than liquidation when the belief about the occurrence of the two states coincides with the prior. It implies that without additional information lender(s) perceives continuation as more profitable than liquidation. To ascertain the profitability of liquidation, lender(s) must acquire additional information.

The manager, having no wealth of his own, must seek external financing. There are two types of lenders--banks and public lenders. Unlike public lenders, a bank lender can have access to a costly monitoring technology which generates an interim signal correlated with the realized state. We assume that only the bank lender who lent at t=0 can observe a signal at t=1[3]. The bank's signal is, however, not verifiable, and therefore can not be contracted upon. By expending effort $e\in[0,1]$, the banker can observe the realized interim state with probability e and remain uninformed with probability 1-e[4]. Expending effort e costs the bank $\psi(e)$. We assume $\psi'(e)>0$, $\psi''(e)>0$ and $\psi(e=0)=0$. At t=0, there is competitive supply of public and bank financing. The prevailing interest rate is normalized to zero. This implies that, ex ante, lenders are willing to provide financing if the expected returns from their claims equal the amount lent. At t=1, the supply of public financing remains perfectly competitive. The only type of contract between the firm and its lenders is the standard debt contract which specifies a repayment schedule and contains a covenant. Following the incomplete contract approach, we assume that the interim decisions are not contractible. Therefore, the interim

---

[3]This is consistent with most of the empirical findings. See Billett, Flannery and Garfinkel (1995), James (1987), Lummer and McConnell (1991).

[4]Formally, this corresponds to the information structure in which the signal space consists of three elements, $\{s_H, s_L, \phi\}$. The correlations between the signals and the interim states are $f(s_H|\theta_H)=e, f(\phi|\theta_H)=1-e$, $f(s_L|\theta_L)=e$ and $f(\phi|\theta_L)=1-e$. The signal ϕ is completely uninformative, and the signal s_H (s_L) perfectly reveals the state θ_H (θ_L).

continuation decision can not be explicitly constrained by the covenant, and the party who has the interim control rights can unilaterally decide on the actions to be taken[5]. All agents are assumed to be risk neutral.

At t=0, the manager attempts to raise the capital needed to start the project. Apart from being the residual claimant of the t=2 cash flow, the manager enjoys a non-transferable and state-contingent private control rent $C(\theta)$, with $C(\theta_H)>C(\theta_L)$, provided that the project is continued to t=2[6]. The private control rent is the source of potential misalignment of incentives, between the manager and its lenders, over the interim continuation decision. To raise initial financing, the manager must credibly assure lenders of an expected repayment equal to the funds they initially provide. We assume that the return from the project also satisfies the following condition.

Assumption 4.2: $$\nu_H \int_0^X r dF(r|\theta_H)+\nu_L L > I_0 > E_\theta[\int_0^X r dF(r|\theta)]. \quad (4\text{-}3)$$

Assumption 4.2 indicates that the t=0 expected cash flow is less than the initial investment if the project is always continued to t=2; It exceeds the initial investment if the project is only continued in state θ_H. For lenders to break even, liquidation in state θ_L must occur with strictly positive probability and it must strictly benefit the lenders. To simplify our notations, we introduce the following definition.

---

[5]This arises either when an informed bank's signal is not verifiable or if the costs of describing interim actions are prohibitively high.

[6]The control rent may not be actual monetary benefit for the manager. In our analysis, it merely serves as a measure of the divergence of preference, between lenders and the manager, over the interim continuation decision. For further discussions, see Aghion and Bolton (1991). Hart and Moore (1991) endogenize the control rent by giving the manager some bargaining power through his ability to quit.

Definition 4.1: $$r(\theta)=\int_0^X rdF(r|\theta), \qquad \theta\in\{\theta_H,\theta_L\}. \tag{4-4}$$

The information structure is specified as follows. At t=0, information is symmetric among all agents, with common prior of the favorable and the unfavorable state being v_H and v_L respectively. The manager can costlessly observe the interim state and the signal acquired by the bank. Thus, at t=1, the manager knows whether the bank is informed or not. On the other hand, public debtors observe neither the realized state nor the signal acquired by the bank.

If the project is financed at t=0, the firm and the bank can costlessly renegotiate at t=1. Renegotiations occur either because there are needs to modify the terms of the bank loan or because the firm needs to request additional financing. The renegotiation process is specified as follows. We assume that both the bank and the firm can initiate renegotiation. In bargaining over a new contract, the firm can, with probability λ, make a take-it-or-leave-it offer which the bank can accept or reject; With probability $1-\lambda$, the bank can make a take-it-or-leave-it offer which the firm can accept or reject. Thus, λ is a measure of the firm's bargaining power[7]. In the event that renegotiation breaks down, two possibilities arise: 1)If the firm has short-term bank debt outstanding, then the bank can either force liquidation or forgive the short-term repayment and allow the project to continue (if possible)[8]; 2) If the firm

[7]There are many factors which can affect the size of λ. For example, the bank's reputational concerns, the length of the firm-bank relationship, and the firms' accessibility to alternative sources of capital, can all affect the relative bargaining power.

[8]Since this assumption will turn out to be rather important for our analysis, we provide some justifications. First, banks are not prohibited by law from forgiving repayments due. Second, as is easily seen , if the bank only holds short-term debt claim, it will not forgive the t=1 repayment if renegotiation breaks down. However, when it holds both short-term and long-term claims, it may choose to forgive the t=1 repayment in the interest of capturing a larger

only has long-term debt obligation(s) outstanding, the existing contract(s) stands. On the other hand, we assume that it is impossible for the firm to renegotiate with the existing public debtholders[9]. The firm can, however, raise financing from the interim competitive capital market, in which case the firm makes a take-it-or-leave-it offer to investors[10].

As a preliminary analysis, we demonstrate the existence of demand for bank debt. Suppose the firm tries to raise initial financing by only borrowing public debt. With public lenders, monitoring and interim renegotiation are both infeasible. Without monitoring, public lenders can not acquire additional information about the realized state, and, by assumption 4.1, they will not choose to liquidate the project. Since renegotiation is infeasible, the firm will never choose to liquidate the project. Thus, the project is always continued, and by assumption 4.2 public debtors will not finance the project. To raise initial financing , the firm's initial borrowing must include bank debt.

When the firm's initial borrowing includes bank debt, the lender's (or lenders') t=0 expected payoff can be written as

---

t=2 payoff. This is similar to a debt restructuring except that the latter is usually furnished under a new contract. The difference arises because in the present setting, financial distress occurs with certainty. Therefore, the result of a t=1 restructuring is partially reflected in the ex ante contract. Finally, this assumption accords our definition of the interim control rights to that of Grossman and Hart (1986). According to these authors, the party who has the control rights can unilaterally decide on the course of action in the absence of negotiation. In our model, the there are two possible interim actions--continuation and liquidation. By allowing the bank to forgive the t=1 repayment, short-term bank debt claim confers the control rights on the bank in the sense of Grossman and Hart (1986)

[9]This assumption may be justified by the presence of free rider problem in exchange offers. See Gertner and Scharfstein (1994).

[10]We assume that lenders in the capital market are sufficiently diverse so that the firm can make the offer to investors other than the existing ones.

$$(1-e)R_u+eR_i-\psi(e)=R_u+e(R_i-R_u)-\psi(e), \tag{4-5}$$

where R_i and R_u are the lender's (or lenders') payoff when the bank is informed and uninformed respectively, and e is the bank's monitoring effort. To raise initial financing, the expected payoff, in (4-5), must be at least as large as the initial investment, I_0. The analysis in this paper is based on the observation that there are deadweight costs associated with monitoring[11]: 1) There is a cost $\psi(e)$ for expending monitoring effort e; And 2) as a result of monitoring, the project must be liquidated when the bank is informed of the unfavorable state. Liquidation destroys the manager's private rent, which may exceed the liquidation value of the assets (see the discussion in section III). Thus, the manager desires to minimize the level of monitoring required by initial financing. To minimize the required level of monitoring, the manager must maximize the lender's (or lenders') payoff when the bank is uninformed, R_u, and, when positive level of monitoring is required by initial financing, the lender's (or lenders') marginal benefit of monitoring, $R_i - R_u$. Furthermore, he must structure the bank's debt claim to induce the minimum level of monitoring.

The Mix of Long-Term Public and Bank Debt Claims

The manager raises initial financing by borrowing from a bank and public lenders. In return, he promises a t=2 repayment s_2 to the bank and t_2 to public lenders. The bank's and the firm's expected payoff in state θ is denoted as $R_b(s_2,t_2|\theta)$ and $R_f(s_2,t_2|\theta)$ respectively[12].

---

[11]We assume throughout the paper that all rents extracted by the lenders in excess to the initial funds lend are prepaid. Thus, the manager's objective is to maximize the total t=0 expected surplus, while ensuring initial financing.

[12]Both of the payoffs incorporate any contractual specifications which may affect them, including, for example, the relative seniority between the public and bank debt claims.

The mix also specifies the payoffs to the bank and to the public debtors in liquidation, denoted as L_b^t and L_p^t respectively. In analyzing the mixes of long-term debt claims, we make

Assumption 4.3: $C_L \leq L - r(\theta_L)$[13]. (4-6)

Assumption 4.3 implies that liquidation in state θ_L generates positive surplus, and hence is efficient. Since the manager can not renegotiate with the public lenders, it falls on the bank to bribe the manager and induce liquidation. Consider the interim renegotiation when the bank is informed. In this case, the firm and the bank negotiate under symmetric information and the their total payoff is maximized[14]. If the project is continued, the firm's and the bank's individual rationality conditions imply that the existing contract will not be replaced. If the project is liquidated, the total cash flow available to the firm and the bank is L_b^t. Thus, negotiation leads to liquidation in state θ_L if and only if

$$L_b^t \geq [R_f(s_2, t_2 | \theta_L) + R_b(s_2, t_2 | \theta_L) + C_L]. \quad (4\text{-}7)$$

Condition (4-7) indicates that it is individually rational for the bank to bribe the manager and induce liquidation in state θ_L. When the bank is uninformed, the analysis is more involved and

---

[13]If $C_L > L - r(\theta_L)$, then the project can not be financed by a mix of long-term debt. This follows because the manager can always continue the project without raising additional financing at t=1. To induce liquidation in state θ_L, lenders must offer the manager a bribe of at least C_L. This leaves the lenders (both the bank lender and the public debtors) with a t=0 expected payoff no greater than

$$\nu_H r(\theta_H) + \nu_L [L - C_L] \leq E_\theta [r(\theta)].$$

By assumption 4.2, this is smaller than the initial investment I_0. Therefore lenders will refuse to provide initial financing.

[14]In general, when one of the parties in the negotiation is wealth constrained, symmetric information is not sufficient to ensure the optimality of bilateral bargaining. In the present case, monetary transfer is from the bank, who is not subject to wealth constraint, to the manager, and therefore optimality is ensured.

is relegated to the appendix. We summarize the result in the following lemma.

Lemma 4.1:If the bank is uninformed, then in the unique equilibrium of interim negotiation[15]:

i)If s_2 and t_2 satisfy (4-7) and

$$L_b^1-[R_f(s_2,t_2|\theta_H)+C_H]\leq E_\theta[R_b(s_2,t_2|\theta)], \tag{4-8}$$

then the project is liquidated in state θ_L and continued in state θ_H, independent of who makes the offer;

ii)For all s_2 and t_2 which do not satisfy (4-7) or (4-8), the project is either always liquidated or always continued when the firm makes an offer.

In the following, we focus on debt structures which satisfy conditions (4-7) and (4-8). Condition (4-8) indicates that when the firm makes an offer in state θ_H, continuing the project returns it a larger payoff than that in liquidation. It ensures that the project is always continued in the favorable state. In (4-8), the bank's payoff in continuation is its expected return, reflecting that the bank is uninformed. If the debt structure satisfies conditions (4-7) and (4-8), then, independent of the bank's information, the equilibrium for the interim negotiation is separating: The project is liquidated in state θ_L and continued in state θ_H. The manager desires to maximize the total surplus while ensuring initial financing. Given the above liquidation policy, the total surplus is independent of the promised repayments to the debtholders. Thus, to ensure initial financing, the manager desires to maximize the lenders'

[15]It is well known that negotiation under asymmetric information generally leads to multiple equilibria when the informed party can also propose offers. Throughout this analysis, we require that the equilibrium satisfy the divinity criterion of Banks and Sobel (1987).

(the sum of the bank's and the public debtor's profit) expected payoff[16]. The following proposition characterizes the optimal solution.

Proposition 4.1: At optimal, $s_2+t_2=X$ and

$$L_b^1=R_b(s_2,t_2|\theta_L)+C_L. \tag{4-9}$$

The project is liquidated in state θ_L and continued in state θ_H, and the bank does not monitor.

In the optimal solution, the bank's monitoring effort is zero. Equation (4-5) implies that the lenders' t=0 expected payoff consists of only R_u. Since the total promised repayment to the lenders, s_2+t_2, equals the maximum profit from the project, they acquire the entire cash flow from the project in state θ_H. To induce liquidation in state θ_L, however, the bank must offer the manager a bribe C_L. The lenders'--the bank and the public lenders--t=0 expected profit is

$$R_u=v_H r(\theta_H)+v_L[L-C_L]. \tag{4-10}$$

Thus, if the payoff, in (4-10), exceeds the initial investment, I_0, project can be financed by a mix of long-term public and bank debt.

The optimal solution has the following two properties. First, since the continuation decision is independent of the bank's information, there is no need for monitoring. There are two reasons why the interim equilibrium is separating even when the bank is uninformed: 1)Since the manager's control rent is sufficiently small, he can recoup the loss of private rent when the project is liquidated; 2)Since the manager has the interim control rights, he is not harmed by revealing the unfavorable state.

---

[16]From part ii) of lemma 4.1, debt structures which do not satisfy (4-6) or (4-7) reduces both the total surplus and the contractible cash flow. Consequently, they are suboptimal.

Next, consider condition (4-9) in proposition 4.1. The left hand side of the equality is the total payoff to the bank and the firm when the project is liquidated in state θ_L. The right hand side is their total payoff if the project is continued. Condition (4-9) indicates that the bank and the firm do not strictly benefit from liquidation; The public debtors extract the entire surplus from liquidation. As has been pointed out previously, the manager desires to maximize the t=0 expected payoff to the lenders. Since the bank's debt claim is not renegotiated when the project is continued, the firm can credibly pay out the profit from continuation through ex ante contracting. However, when the project is liquidated, the bank's debt claim must be renegotiated. If the bank is the only lender, the division of the surplus from liquidation can not be specified through ex ante contracting. Instead, it is divided according to the relative bargaining power. When the manager commands large bargaining power, he can extract most of the surplus from liquidation. The bank's benefit from liquidation diminishes and, by assumption 4.2, it may refuse to provide initial financing. Unlike the bank lender, however, public debtors can free ride on the benefit of the negotiation between the bank and the firm without making concessions of their claims in liquidation. Thus, by borrowing a mix of bank debt and public debt, the firm can credibly pay out the surplus generated from liquidation to the public lenders. The optimal debt structure maximizes debtors' total payoff when the bank is uninformed, R_u, and initial financing is ensured without the firm incurring any cost of monitoring. This provides an explanation why the firm may desire to diversify its borrowing[17].

---

[17]Raja (1992) argues that bank's opportunism may cause public debt to be more desirable than bank debt. Our finding extends Raja's result in three aspects: 1) It indicates that a mixed debt structure can be desirable; 2) Diversified borrowing also arises when the firm commands large

Optimal Bank Debt

When the manager's control rent becomes large, that is, when the divergence of preference between the manger and the lenders is large, liquidation through interim renegotiation is no longer feasible. Therefore, the firm can not finance the project by using a mix of long-term public and bank debt. Specifically, we will assume the following in the ensuing analysis.

Assumption 4.4: $C_L > L$. (4-11)

In this case, feasibility of initial financing requires involuntary liquidation. This implies that the firm's initial borrowing must include short-term debt. A short-term debt claim transfers the t=1 control rights to the lender when the manager can not make a repayment. Given the control rights, the lender can force liquidation without having to bribe the manager.

By assumption 4.1, without additional information lenders' consider continuation as more profitable than liquidation, and the firm can always continue the project by promising a sufficiently large t=2 repayment. Since initial financing requires that the project be liquidated in the unfavorable state with strictly positive probability, the firm must structure the debt claim to induce monitoring by the bank lender. To find the minimum level of monitoring required by initial financing, notice that the maximum value for R_u, the lender's (or lenders') expected payoff when the bank is uninformed, is $E_\theta[r(\theta)]$ and the maximum value for R_i, the lender's (or lenders') expected payoff when the bank is uninformed, is $\nu_H r(\theta_H)+\nu_L L$. The minimum level of monitoring required by initial financing, e^*, is then

---

bargaining power; 3)Diversified borrowing may be desirable even in the absence of moral hazard problem.

defined by

$$E_{\theta}[r(\theta)]+v_L e^*[L-r(\theta_L)]-\psi(e^*)=I_0. \tag{4-12}$$

In the ensuing analysis, we make

Assumption 4.5:
$$\frac{d\psi(e^*)}{de^*}\leq v_L[L-r(\theta_L)]. \tag{4-13}$$

To interpret assumption 4.5, suppose that the firm can structure the debt claims so that both R_u and R_l attain their maximum values. The right hand side of (4-13) is the lender's (or lenders') total marginal benefit of monitoring which must (weakly) exceed the bank's marginal benefit from monitoring. Assumption 4.5 ensures that it is feasible for the firm to both minimize and induce the amount of monitoring required by initial financing.

In this section, we assume that, ex ante, the firm only borrows from a bank. At t=1, however, it can acquire additional financing from the market and/or by negotiating with the bank. Besides motivating the analysis in the next section, this section derives the optimal debt structure for a firm which initially does not have access to the public debt market.

Short-Term Bank Debt

The manager raises initial financing by borrowing from a bank and promises to repay s_1 at t=1. In the discussion of short-term bank debt, we assume that the firm does not raise financing from the interim market. To continue the project, the manager must renegotiate with the bank. If the renegotiation breaks down, the project will be liquidated and the bank gets L while the manager gets nothing. These are the bank's and the firm's reservation levels in the renegotiation. Consider the renegotiation when the bank is informed. In state θ_H, when

the bank makes an offer, it can demand the entire return from the project by proposing a continuation contract $s_2^b=X$[18]. When the firm makes an offer, it promises a t=2 repayment which returns the bank an expected payoff equal to its reservation level, L. In state θ_L, the project must be liquidated. When the bank is uninformed, the analysis is slightly more complicated and is relegated to the appendix. The result is summarized in the following lemma.

Lemma 4.2: When the bank is uninformed, there is a unique equilibrium in the interim renegotiation. In this equilibrium, the project is always continued and

i) the firm proposes the pooling continuation contract s_2^f which yields the bank an expected return equal to its reservation level L;

ii)the bank proposes the pooling continuation contract $s_2^b=X$.

The intuition for lemma 4.2 is quite simple. With short-term debt claim, the bank can unilaterally decide to terminate the project. If the manager's offer is state contingent, the bank can infer each realized state. It will refuse the offer indicating the unfavorable state and terminate the project. The manager thus offers the same contract in both states, and the feasibility of continuation follows from assumption 1. When the bank makes an offer, it can demand the entire t=2 cash flow by proposing a continuation contract $s_2^b=X$. Alternatively,

---

[18]In the negotiation between the firm and the bank, the parties can reach a new contract which specifies a t=1 repayment to the bank. The project is then liquidated and the bank is paid according to the new contract. Such a contract is termed a liquidation contract. Alternatively, the parties can reach a new contract which specifies a t=2 repayment. The project is then allowed to continue. Such a contract is termed a continuation contract. Throughout the paper, offers with subscript 1 indicate t=1 repayments and correspond to liquidation contracts. Offers with subscript 2 indicate t=2 repayments and correspond to continuation contracts. All offers are described in terms of the payments to the bank.

the bank can offer a menu of contracts which screens out the unfavorable state and induce liquidation. However, the manager will reveal the unfavorable state only if he is bribed, at least, his private rent C_L. By assumption 4.3, this is infeasible. Part ii) of lemma 4.2 then follows.

It follows from the above discussion that the bank's t=0 expected profit and monitoring effort are

$$P_b^0(\lambda)=(1-\lambda)E_\theta[r(\theta)]+\lambda L+e(1-\lambda)v_L[L-r(\theta_L)]-\psi(e), \tag{4-14}$$

$$\frac{d\psi(e)}{de}=(1-\lambda)v_L[L-r(\theta_L)]. \tag{4-15}$$

Equation (4-14) indicates that the bank's profit is independent of the manager's private rent. This follows because the short-term debt claim transfers the interim control rights to the bank, allowing it to force liquidation without bribing the manager. However, the bank's expected payoff depends on its bargaining power, $1-\lambda$. At t=1, the bank acquires the control of the firm's assets. To continue the project, the manager must purchase the assets back from the bank through renegotiation[19]. The expected price he must pay increases with the bank's bargaining power. Consequently, the bank's expected payoff increases with its bargaining power.

Equation (4-15) indicates that the bank's marginal profit of monitoring is positive and is increasing in the bank's bargain power, $1-\lambda$. When the bank only holds long-term debt claim, its marginal profit of monitoring is zero because it can not act upon better information

[19] Hart and Moore (1989, 1995) show that the need to purchase control rights from the debtors can serve to discipline the manager by forcing him to pay out excess cash, thereby mitigating the free cash flow problem suggested by Jensen (1986).

without the control rights. Given the control rights, however, the bank can benefit from better information which allows it to timely liquidate the project in the unfavorable state. Thus, maximizing the value of the control rights motivates the bank to monitor[20].

To assess the feasibility and desirability of short-term bank debt financing, let λ^* be defined as $P_b^0(\lambda^*)=I_0$. From assumption 4.5, $0 \le \lambda^* < 1$ and $\lambda^*=0$ only when (4-13) holds with equality. If the firm's bargaining power is sufficiently large, so that $\lambda > \lambda^*$, short-term bank debt financing is infeasible. In this case, the manager can repossess the assets at a small average price, whenever the project is allowed to continue. He extracts most of the surplus from continuation, and the bank can not recoup the initial fund lent. On the other hand, if the firm's bargaining power is sufficiently small, so that $\lambda \le \lambda^*$, short-term bank debt financing is feasible. However, except when $\lambda^*=0$, the monitoring effort induced strictly exceeds the minimum level of monitoring e^*. This follows because when the firm raises initial financing by using only bank debt, the bank acquires the entire benefit from monitoring and the monitoring effort supplied exceeds e^*. Thus, the firm's desire to minimize the required level of monitoring through maximizing the marginal benefit of monitoring, and its desire to induce the minimum level of monitoring through controlling the bank's benefit from monitoring are in conflict with each other.

Bank Debt Requiring Both Short-Term and Long-Term Repayments

The manager raises initial financing by borrowing from a bank. In return, he promises to repay $s_1>0$ at t=1 and s_2 at t=2. At t=1, the manager can finance the repayment, s_1, by

---

[20]Park(1995), Raja and Winton (1995) show that an alternative way to confer control right upon the bank is to combine covenant with long term bank loan.

raising funds from the public debt market and/or by negotiating with the bank. To raise financing from the market, the manager makes a take-it-or-leave-it offer to investors. This offer consists of the amount of money the manager intends to borrow and, in exchange, the debt claim. Since the public debt market remains competitive at t=1, this debt claim is determined by investors' individual rationality conditions. It follows that at interim the manager's strategic decision involves choosing the amount of financing to be raised from the market. For simplicity, we assume that the firm's offer is observable by the bank, while public lenders do not observe the outcome of the negotiation between the firm and the bank. The timing of the interim game is specified as follows. First, the firm decides whether or not to raise financing from the market, and the amount of financing to be raised. After acquiring the fund, the firm makes a repayment to the bank[21]. If the bank is not fully repaid, then the manager must still negotiate with the bank. The following lemma summarizes the manager's equilibrium strategy.

Lemma 4.3: In the interim equilibrium induced by the optimal bank debt,

i)if $L>R_b(s_2|\theta_H)$ and the firm raises financing from the interim market, then its equilibrium offers must be separating;

ii) if $L\leq R_b(s_2|\theta_H)$, then the firm never raises financing from the interim market.

Part i) of lemma 4.3 indicates that if the firm raises financing from the interim market, its equilibrium offers must be separating: Its offer when the bank is informed of the favorable

[21]Here, we make two assumptions. First, we assume that in equilibrium the firm will not make an offer to investors if the offer will be rejected. This assumption will hold if there is flotation costs in issuing public debt. Second, we assume that the manager can carry the funds to t=2, but he can not divert the money to gain private benefit.

state must be different from that when the bank is uninformed. This follows from the following two reasons. First, it is strictly suboptimal for the firm to renegotiate with the bank after it has made a repayment, because borrowing from public lenders reduces the total surplus over which the firm and the bank bargain. The firm is strictly better off directly negotiate with the bank without making a repayment. Thus, if the firm raises financing from the interim market, it will borrow a sufficient amount so that after the repayment the bank will forgive the residual t=1 repayment. On the other hand, since the firm is the residual claimant of the t=2 cash flow, it also desires to minimize the promised t=2 repayment to public lenders, and therefore the amount of financing to be raised. These two considerations suggest that the firm desires to raise an amount of financing from the market so that after the repayment, the bank is just willing to forgive the residual t=1 repayment and allow the project to continue. Second, the t=2 claim specified in the initial contract is more valuable to the bank when it is informed of the favorable state than when it is uninformed. Therefore, it will allow the project to continue with a smaller repayment in the former case. These two reasons suggest that the interim equilibrium must be separating. Note that this feature of the equilibrium indicates that at interim the market can perfectly infer the bank's information. In other words, the bank's information is transmitted to the market[22].

Part ii) of lemma 4.3 indicates that when the promised t=2 repayment is sufficiently

---

[22]The bank's decision to forgive the t=1 residual repayment can also be interpreted as that the bank automatically provides refinancing in an amount equal to the residual repayment due. According to this interpretation, for a fixed t=2 repayment, the bank provides a larger amount of refinancing when it is informed of favorable state. In other words, the bank charges a lower interest for the interim refinancing when it is informed of the favorable state. Thus, in equilibrium investors revise their belief upward when the bank charges a lower interest for refinancing.

large, the manager never desires to raise financing from the interim market. To see this, note that, if the negotiation breaks down when the bank is informed of the state θ_H, it can get a payoff L by liquidating the firm or a payoff $R_b(s_2|\theta_H)$ by forgiving s_1 and allowing the project to continue. If $L \leq R_b(s_2|\theta_H)$, the bank prefers to let the project continue when the negotiation fails. Therefore, if the manager does not raise financing from the market and forces the negotiation to break down[23], the bank will not liquidate the project. Anticipating that continuation is ensured without making a repayment, the manager will not raise financing from the interim market. When the bank is uninformed, the same reasoning indicates that the manager will not raise interim financing from the market if $L \leq E_\theta[R_b(s_2|\theta)]$. Consider now the case when

$$E_\theta[R_b(s_2|\theta)] < L \leq R_b(s_2|\theta_H). \tag{4-16}$$

Since $E_\theta[R_b(s_2|\theta)] < L$, if the negotiation breaks down the bank will not forgive the t=1 repayment when it is uninformed. However, in the interim equilibrium induced by the optimal bank debt claim, the firm will not raise financing from the market. To see this, notice that an increase in the bank's payoff, when it is uninformed, reduces its marginal benefit from monitoring without decreasing its t=0 expected payoff. Thus, to minimize the required level of monitoring, the manager desires to structure the debt claim to constraint himself from raising financing from the market when the bank is uninformed[24]. Consequently, in the equilibrium induced by the optimal bank debt claim, the manager does not raise financing from

---

[23]The manager can force the negotiation to break down by, for example, rejecting any of the bank's offer and offer a t=2 repayment of zero.

[24]This can be easily ensured by setting a sufficiently large t=1 repayment senior.

the market. The project is then continued through negotiation.

Before describing the optimal bank debt, we explain figure 4-1. In the figure, s_2^u is the promised t=2 repayment which returns the bank, when it is uninformed, an expected payoff equal to the liquidation value of the assets, L. If the firm's promised t=2 repayment exceeds s_2^u, the uninformed bank will forgive the t=1 repayment if renegotiation breaks down. Otherwise, it will choose to liquidate the project. Similar interpretation applies to s_2^H when the bank is informed of the favorable state. Since $s_2^u(\lambda) \geq s_2^u$, the uninformed bank will forgive the t=1 repayment if the promised t=2 repayment is equal to $s_2^u(\lambda)$. The expected payoff it will receive is the same as that when the project is continued through renegotiation. Similar interpretation applies to $s_2^H(\lambda)$ when the bank is informed of the state θ_H. From the figure, it is clear that $s_2^u(\lambda=1)=s_2^u$ and $s_2^H(\lambda=1)=s_2^H$.

Consider first the case when the manager has large bargaining power so that $\lambda>\lambda^*$, and financing by short-term bank debt is infeasible. We denote the bank's expected payoff, in state θ, from a promised t=2 repayment x as $R_b(x|\theta_H)$. As will soon become clear, we only need to focus on structures in which $L \leq R_b(s_2|\theta_H)$, where s_2 is the firm's promised t=2 repayment in the initial contract. The following proposition characterizes the optimal bank debt when the firm commands large bargaining power.

Proposition 4.2: i)If in the optimal bank debt $s_2 \geq s_2^u$, then $s_2=X$ and the t=0 bank's expected payoff is $P_b^0(\lambda=1)$, where $P_b^0(\lambda)$ is defined in (4-14);

ii)If in the optimal bank debt $s_2^H(\lambda)<s_2<s_2^u$, then the t=0 bank's expected payoff and monitoring effort are

$$\lambda L+(1-\lambda)E_\theta[R_b(X|\theta)]+e(s_2)v_L(1-\lambda)[L-R_b(X|\theta_L)+A(s_2)]-\psi(e(s_2)), \qquad (4\text{-}17)$$

$$\frac{d\psi(e(s_2))}{de(s_2)} = v_L(1-\lambda)[L - R_b(x|\theta_L)] + A(s_2). \tag{4-18}$$

$A(s_2)$ is continuous and strictly increasing in s_2 with $A(s_2=s_2^H(\lambda))=0$.

From fig.4-1, if $s_2 \geq s_2^u$, the bank will forgive the t=1 repayment both when it is informed of the favorable state and when it is uninformed. Therefore, its expected payoff in continuation is completely specified by the initial contract. Clearly, the bank's t=0 expected payoff is increasing in s_2. Its monitoring effort, however, is decreasing in s_2. This follows because the bank's payoff in continuation is independent of its information when the favorable state is realized. In the unfavorable state, the bank's loss of profit from allowing the project to continue decreases with s_2. Thus, the benefit from timely liquidation diminishes and the bank's monitoring effort decreases as s_2 increases. Since the manager desires to minimize monitoring as long as he can raise initial financing, he chooses to set s_2 at its maximum. In this case, the bank acquires all the cash flow from the project when it is continued. Comparing with the case of short-term debt claim, the bank is effectively assuming full interim bargaining power, and initial financing becomes feasible.

Debt claim described in part ii) of proposition 4.2 becomes optimal when $s_2^H(\lambda)<s_2^u$ and λ is just above λ^*. Under these conditions, the bank's t=0 expected payoff from a short-term claim is just below the required initial investment. Thus, with a slight increase in the bank's expected payoff initial financing becomes feasible. If $s_2^H(\lambda)<s_2^u$, the firm can increase the bank's expected payoff by promising a t=2 repayment s_2, so that $s_2^H(\lambda)<s_2<s_2^u$. From fig.4-1, comparing with the short-term claim, this contract returns the bank a higher payoff when it is informed of state θ_H; When the bank is uninformed, it returns the bank the same payoff. By continuously increasing s_2 over the interval $[s_2^H(\lambda), s_2^u)$,

both the bank's t=0 expected payoff and its monitoring effort are continuously increased. When initial financing only requires a small increase in the bank's t=0 expected payoff, the contract in part ii) of proposition 4.2 becomes optimal, because while ensuring initial financing it reduces the amount of monitoring.

Consider next the case when the bank has large bargaining power, so that $\lambda<\lambda^*$. In this case, financing by short-term bank debt is undesirable because of the bank's oversupply of monitoring. The following proposition characterizes the optimal bank debt claim.

Proposition 4.3: With the optimal debt claim, the bank's t=0 expected payoff equals the initial investment I_0. At interim, when the bank is uninformed, its payoff is the same as that with short-term claim. This outcome can be implemented by a repayment schedule in which the t=1 promised repayment is senior and the t=2 repayment is arbitrarily small. In the interim equilibrium induced by this structure, the firm seeks financing from the interim market only when the bank is informed of the state θ_H.

Form part ii) of lemma 4.3, the firm can not increase the total surplus for the bank and itself by raising financing from the interim market, because the firm is fairly priced by the market. It follows that, independent of whether the firm raises financing from the market or not, the uninformed bank's payoff can not exceed that with short-term claim. To induce the desired level of monitoring, the firm must reduce the bank's payoff when it is informed of the favorable state. This, as is indicated by proposition 4.3, can be implemented by a senior, and "almost" short-term, bank debt claim. As is shown in the appendix, in the interim equilibrium, the firm raises financing from the market when the bank is informed of the favorable state, so that after the repayment the bank is just indifferent between liquidating the firm and forgiving

the residual short-term repayment. Thus, when the bank is informed of the favorable state, its payoff is equal to the promised t=1 repayment. By optimally setting the value of this repayment, the firm can induce the desired level of monitoring.

The discussion in this section indicates that, in general, the firm prefers to borrow bank debt requiring repayments both at interim and on the final date. Such a repayment schedule gives the firm more latitude in achieving its two goals in designing the optimal debt structure. Optimal bank debt generates two types of outcome. In the first instance, as indicated by part I) of proposition 4.2, both the bank's payoff when it is uninformed, R_u and its marginal benefit from monitoring, R_b, are maximized. Therefore, the required level of monitoring for initial financing is minimized. However, the firm can not further adjust the level of monitoring induced and there is an over supply of monitoring effort. In the second instance, as indicated in part ii) of proposition 4.2 and proposition 4.3, the firm can adjust the level of monitoring supplied, but neither R_u nor R_b is maximized. Consequently, the amount of monitoring required by initial financing is not minimized. Thus, if the firm raises initial financing by only borrowing from a bank, it can not simultaneously achieve both of its two goals.

The Optimal Mix

The analysis in the previous section assumes that the manager raises initial financing only from a bank. Returning to the discussion of the optimal mix, we will analyze in this section the two cases in which the firm acquires initial financing by borrowing, in addition to bank debt, either short-term or long-term public debt. For ease of exposition, however, we will assume away interim market. A complete analysis, which takes into considerations of the

interim public debt market, proceeds in the same way as in the previous section.

Mixed structure with bank debt and long-term public debt

The firm raises initial financing by borrowing from a bank and public lenders. In return, it promises the bank a repayment $s_1>0$ at t=1 and s_2 at t=2, and the public lenders a repayment t_2 at t=2. The bank's and the public lenders' payoff in state θ from the promised t=2 repayments s_2 and t_2 are denoted as $R_b(s_2,t_2|\theta)$ and $R_p(s_2,t_2|\theta)$[25]. In addition, the mixed structure specifies the bank's and the public lenders' payoffs in liquidation. We denote the bank's payoff in liquidation as L_b. Lengthy but straightforward calculations provide the following result.

Lemma 4.4: In the optimal mix, $L_b \le E_\theta[R_b(s_2,t_2|\theta)]$.

Lemma 4.4 implies that at interim the bank will forgive the t=1 repayment both when it is informed of the favorable state and when it is uninformed. Thus, the bank's monitoring effort and the debtor's total t=0 expected profit (the bank's and the public debtors') are

$$\frac{d\psi(e)}{de} = v_L[L_b - R_b(s_2,t_2|\theta_L)], \tag{4-19}$$

$$(1-e)E_\theta[R_b(s_2,t_2|\theta)+R_p(s_2,t_2|\theta)]+e[v_H(R_b(s_2,t_2|\theta_H)+R_p(s_2,t_2|\theta_H))+v_L L]-\psi(e). \tag{4-20}$$

Equations (4-19) and (4-20) imply the following. First, a comparison between (4-5) and (4-20) indicates that both R_u, the creditors' payoff when the bank is informed, and R, the creditors' payoff when the bank is uninformed, depend only on the sum s_2+t_2. Second, the debtors' marginal benefit from timely liquidation, or from monitoring, can be written as

[25] Again, these payoffs incorporate any contractual specifications which may affect them, including, for example, the relative seniority between the public and bank debt claims.

$$v_L[L_b - R_b(s_2,t_2|\theta_L)] + v_L[L - L_b - R_p(s_2,t_2|\theta_L)]. \quad (4\text{-}21)$$

Equation (4-21) reveals that, in contrast to the case when the firm raises initial financing only from a bank, the debtors' total marginal benefit of monitoring is no longer the same as the bank's marginal benefit from monitoring. The public debt claim breaks this linkage, allowing the manager to separate the tasks of maximizing the lenders' total benefit of monitoring and inducing the bank to supply the desired level of monitoring. To minimize the required level of monitoring, the manager desires to maximize the sum s_2+t_2. Inducing the desired level of monitoring requires that the manager optimally distribute the benefit of monitoring between the bank and the public lenders. The next proposition characterizes the optimal mix.

Proposition 4.4: In the optimal mix,

i)$s_2+t_2=X$;

ii)The bank's monitoring effort is e*;

iii)$L-L_b-R_p(s_2,t_2|\theta_L)\geq 0$;

iv)In the optimal mix, if the payoffs in liquidation are determined by seniority then the public debt claim t_2 is senior to the t=2 bank debt claim s_2.

Part i) and ii) of proposition 4.4 indicates that with the optimal mix of claims, the firm can both minimize and induce the required level of monitoring for initial financing. The public debt claim plays two roles. First, for any fixed payoff scheme for the bank, and so for any fixed level of monitoring, it allows the firm to minimize the required amount of financing by paying out the entire cash flow from the project. Second, by giving the public debtors a share of the proceeds from liquidation, the firm can control the bank's benefit from timely liquidation, and therefore the monitoring effort supplied. Thus, with two debt instruments,

the firm can regulate the bank's incentive to monitor without interfering its desire to minimize the required level of monitoring for initial financing.

Part ii) of proposition 4.4 asserts that when the project is financed by the optimal mix, the public lenders benefit from the bank's monitoring. Therefore, their incentive over monitoring, is aligned with that of the bank's. To see this, notice that the change in the debtors' t=0 expected payoff, under a small change in the bank's monitoring effort, is

$$\delta e[L - L_b - R_b(s_2, t_2 | \theta_L)]. \tag{4-22}$$

If the public lenders' marginal benefit is negative, then a decrease in the bank's monitoring effort increases debtor's t=0 expected payoff. It follows that the firm is strictly better off reducing the the amount of monitoring. Consequently, at optimal, the public lenders' benefit of monitoring must be positive. In fact, the firm prefers to maximize the public lenders' share of the benefit of monitoring, provided the bank supplies the level of monitoring required by initial financing. This follows because while allocating the benefit to the bank can equally increase the debtors' total benefit from monitoring, the firm must incur an increased cost from the increased monitoring by the bank. By allocating this benefit to the public lenders, the firm can increase the debtors' total benefit of monitoring without incurring any additional cost. Thus, when the project is financed by the optimal mix, the bank acts as a delegated monitor.

The reason for part iii) of proposition 4.4 is simple. Lemma 4.4 implies that the promised t=2 repayment to the bank must be strictly larger than its payoff in liquidation. When s_2 is senior to t_2, the bank's payoff in liquidation must be strictly larger than L_b, if $L_b<L$. It follows that when s_2 is senior, the bank's payoff in liquidation is L. Thus, the bank acquires all the proceeds from liquidation, and the manager can not optimally allocate the marginal

benefit of monitoring between the bank and public lenders. Such a mix is therefore suboptimal.

Finally, notice that, in order for the firm to align the public lenders' and the bank's incentive over monitoring and maximize the debtors' total benefit of monitoring, the public debtors' payoff must be state contingent, so that they can benefit from timely liquidation. As we shall see, this is precisely the reason why long-term public debt is strictly preferred to short-term public debt by the firm.

Mixed structure with bank debt and short-term public debt

The firm raises initial financing by borrowing from both a bank and public lenders. In return, it promises to repay the bank $s_1>0$ at t=1 and s_2 at t=2, and the public lenders t_1 at t=1. In addition, the mixed structure defines the payoff in liquidation. We denote the bank's payoff in liquidation as L_b.

At t=1, the firm must renegotiate with the bank. The analysis of the renegotiation is similar to that when the firm only borrows from the bank. The only difference is that the bank's reservation level is increased by t_1, because at interim the bank must invest an additional t_1 to finance the firm's repayment to the public debtors. A complete analysis, however, is not necessary for comparing a mixed structure with short-term public debt and the optimal mix with long-term public debt. We first present the result, followed by an explanation.

Proposition 4.5: A mixed structure with short-term public debt is strictly dominated by the optimal mix with long-term public debt.

Like a long-term debt claim, a short-term public debt claim also allows the firm to

maximize the debtors' payoff when the bank is uninformed, and adjust the level of monitoring induced. By increasing t_1, the firm pledges to pay out more t=2 profit from the project to the bank. This profit is ultimately paid to the public debtors. In addition, by giving the public debtors a share of the proceeds from liquidation, the firm can control the bank's supply of monitoring. Despite the similarities, there is a crucial difference between the short-term and long-term public debt claim. The payoff to the public debtors with long-term claims is state contingent, so that they can benefit from the bank's monitoring. In contrast, the payoff to public lenders with short-term claims is decision dependent. They are fully repaid whenever the project is continued, and they can not benefit from timely liquidations[26]. Therefore, the lenders' total marginal benefit of monitoring is not maximized. It follows that short-term public debt is strictly dominated by its long-term counterpart in the mixed structures.

Empirical Implications

In this section, we list some of the empirical implications of our analysis.

Regulation

Form the discussion in section 3, when agency problems are not sever, i.e. when C_L is small, firms can rely entirely on long term debt. For firms with less discretion over future investment decisions, the agency problems are likely to be small. We would expect that those firms rely more on long-term debt financing. Managers in regulated firms typically have more constrained decision sets comparing to those for managers in the unregulated firms. Our analysis suggests that regulation should increase the average maturity of debt. Barclay and

---

[26]In fact, it can be easily shown that in the optimal mix with short-term public debt, the public lenders' marginal benefit of monitoring is zero.

Smith (1995) find that regulation increases the proportion of the long-term debt by 6.6 percentage points.

Firm size

The analysis in section 4 indicates that debt maturity is intimately related to the relative bargaining power between firms and their banks. Our analysis suggests that firms with large bargaining power over their private lenders must rely more on long-term borrowing. This can effectively enhance the bank's bargaining power and enable initial financing. Taking firm size as a proxy for firms' relative bargaining power, our analysis implies that as firm size increases the maturity of private debt increases. In practice, small firms tend to rely more on short-term bank loans. On the other hand, medium size firms frequently seek financing through private placement, which are usually long-term. Our analysis squares well with this empirical regularity.

Our analysis in section 4 indicates that firms which do not have access to the public debt market may suffer from over-monitoring. Excessive monitoring arises either because a firm can not raise capital from interim public debt market or it can not borrow long-term public debt. It follows that the cost of monitoring is likely to be higher for small firms. This implies that the interest rate of private debt to small firms should, on average, be higher than that to larger firms.

Riskiness

Consider two firms with projects which have the same t=0 expected return and t=1 liquidation value, but differ in their riskiness. Equation (4-21) implies that the firm with the

riskier project demands less monitoring, because the benefit of monitoring increases with the riskiness of the project. It follows that firms with riskier project will finance their project with less bank debt and more public debt. This provides an explanation for the recent empirical finding by Houston and James (1995). In this study, they find that firms with more growth opportunities rely less on bank loan if they borrow only from a single bank. In contrast, for firms with multiple banking relationships, the proportion of bank financing increases. When a firm borrows from a single bank, the bank captures all the benefit from monitoring and it will supply a high level of monitoring. If the firm maintains multiple banking relationships, the benefit of monitoring is shared among the banks and monitoring supplied is lowered. In effect, multiple bank lending creates a quasi-public debt market. Since the cash flows of growth firms are likely to be more volatile than those of more matured firms, our analysis suggests that the change in the proportion of bank debt arises because riskier firms demand less monitoring.

Conclusion

This paper analyzes how firms can optimally design debt structures to facilitate initial financing at minimum costs. Our analysis provides some answers to the questions raised in the introductory section. First, to induce a bank to monitor, it must be given the interim control rights. Given the control rights, the bank can benefit from better information, which allows it to timely liquidate the project in the unfavorable state. The bank's incentive to maximize the value of the control rights motivates it to monitor. Second, the need to diversify a firm's borrowing arises both when the manager's private rent is sufficiently small, so that in the unfavorable state the project can always be liquidated through renegotiation,

and when the manager's private rent is sufficiently large, so that initial financing requires involuntary liquidation. In the first instance, the private debt claim must be renegotiated to allow liquidation. If the firm only borrows from a bank, the division of the surplus from liquidation can not be specified through ex ante contracting. Borrowing public debt allows the firm to credibly pay out the surplus generated from liquidation and facilitate initial financing without incurring any cost of monitoring. In the second instance, the bank must be allocated the interim control rights, so that it can force liquidation without bribing the manager. The control rights also induces monitoring, allowing the bank to ascertain the profitability of liquidation. To minimize the costs associated with monitoring, however, the manager desires to minimize the amount of monitoring required by initial financing. Furthermore, the firm must structure the bank's debt claim to induce the minimum level of monitoring. To minimize the required level of monitoring, the manager desires to maximize the creditors benefit per unit of monitoring effort. To induce the desired level of monitoring, the manager must control the bank's benefit from monitoring. If the firm raises initial financing only from a bank, then the bank acquires all the benefit from monitoring and the manager's two goals are in conflict with each other. Borrowing public debt allows the firm to minimize the required level of monitoring and control the bank's incentive to monitoring through optimally allocating the benefit of monitoring between the bank and the public lenders. Third, the need to control the bank's benefit from monitoring requires that the public lenders be given a share of the benefit from monitoring. Thus, when the project is financed by the optimal mix of long-term public debt and bank debt, the manager prefers to align their incentives over monitoring. The bank thus acts as a delegated monitor.

Finally, we point out some of the limitations of our analysis. First, our analysis ignore the separation of control and ownership: The manager is also the owner of the firm. Strictly speaking, our analysis only applies to small firms in which the agency problems associated with the separation of ownership and control is not sever. Second, our analysis assumes that the intertemporal distribution of the return of the project is fixed. This ignores those cases in which the manager may have discretional choice over the intertemporal distribution of cash flows. Thus, our analysis can not address the problem of how financing by mixed debt structure affects the manager's investment horizon. These issues await future studies.

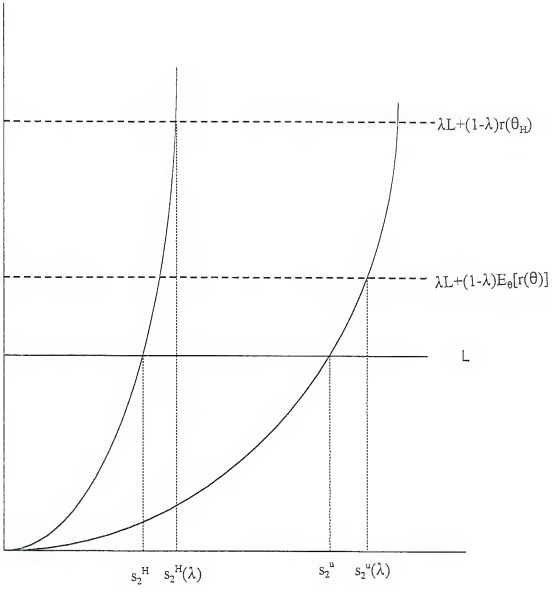


Figure 4-1

CHAPTER 5
CONCLUSIONS

This dissertation has explored three issues in financial economics and law, ranging from optimal compensation schemes for investment advisors to firm's optimal debt structures. The emphasis of this study has been to apply information economics to examine problems the solutions to which depend critically on the allocations of information among agents.

The main findings of this study can be summarized as follows. First, in designing compensation schemes for investment advisors, I show that both the advisors' technologies of information collection and agents' priors are crucial in determining the structures of compensation schemes. The optimal payment scheme rewards the advisor more richly for correctly predicting an outcome, if expending effort best enhances his ability to predict that outcome. When the advisor's information is not publicly observable, I find that the need to induce an advisor to expend effort generally interferes with the need to elicit truthful revelation. Second, in setting optimal legal standards, I find that some divergence between the marginal benefits and marginal costs of providing care by potential violators of the standards are needed to control the costs of enforcement. Furthermore, it is found that setting maximal fines may be welfare reducing. Third, in choosing optimal debt structures, I show that firms in general prefer to borrow from both public and private lenders. Through

monitoring the firm to which they provide funds, private lenders can produce information. In addition, in the optimal debt structure, firms desire to align the public lenders incentive over monitoring with that of private lenders, and so private lenders act as delegated monitors.

While providing interesting findings, this study also reveals that there are important issues that require further analyses. In designing optimal compensation schemes for investment advisors, I have focused on a setting in which there is only one investor and one advisor. Moreover, I have restricted the analysis to a one-period setting. In a fuller analysis, it would be interesting to analyze cases in which one or more of these assumptions are violated. In deriving optimal debt structures, I have ignored agency problems associated with the separation of ownership and control. In addition, in a richer model, one must also relax the assumption that a firm initially borrows only from a single bank. It is also important to examine cases in which firms raise external capital to finance multiple projects. I plan to explore these issues in the future.

APPENDIX A

PROOFS OF THE MAIN RESULTS IN CHAPTER 2

Proof of Proposition 2.1.

The proof of the equivalence between [I-P] and [I-P'] relies on the following two observations. First, the only difference between [I-P] and [I-P'] is in replacing constraints (2-5) and (2-6), in [I-P], by constraints (2-9) and (2-10), in [I-P']. Thus, if it can be shown that a payment scheme satisfies (2-5) and (2-6) iff it satisfies (2-9) and (2-10), i.e, (2-5) and (2-6) are equivalent to (2-9) and (2-10), the equivalence between the two formulations is then established. Second, a payment scheme satisfies (2-5) and (2-6) iff $\{\theta,(x_H,x_L)\}$ is a Nash equilibrium strategy for the advisor. Thus, if it can be shown that a payment scheme satisfies (2-9) and (2-10) iff $\{\theta, (x_H,x_L)\}$ is a Nash equilibrium strategy for the advisor, the equivalence between constraints (2-5),(2-6) and constraints (2-9) and (2-10) is then established. It then follows from the first observation, [I-P] and [I-P'] are equivalent. The following proof shows that a payment scheme satisfies (2-9) and (2-10) iff $\{\theta, (x_H,x_L)\}$ is a Nash equilibrium.

Necessity: If compensation scheme $\{w(x,r): x\in\{x_H,x_L\}\ r\in\{r_H,r_L\}\ \}$ satisfies constraints (2-9) and (2-10), then for any strategy $\{\theta', (x_i,x_j)\}$

$$L(x_i,x_j|\theta')-C(\theta')\leq \mathrm{Max}_{\theta'}[L(x_i,x_j|\theta')-C(\theta')]=\Pi(x_i,x_j)\leq\Pi(x_H,x_L)$$

By (2-10), $\Pi(x_H,x_L)=L(x_H,x_L|\theta)-C(\theta)$. Thus $\{\theta, (x_H,x_L)\}$ is a Nash equilibrium strategy.

Sufficiency: if $\{\theta, (x_H,x_L)\}$ is a Nash equilibrium strategy, then

$$L(x_H,x_L|\theta')-C(\theta')\leq L(x_H,x_L|\theta)-C(\theta) \quad \forall\theta'$$

Hence, $\theta\in Argmax_{\theta'}[L(x_H,x_L|\theta')-C(\theta')]$. Further, for any strategy $\{\theta'_j\ (x_i,x_j)\}$, $x_i,x_j\in\{x_H,x_L\}$,

$$L(x_i,x_j|\theta')-C(\theta')\leq L(x_H,x_L|\theta)-C(\theta)$$

implies $\Pi(x_i,x_j)=Max_{\theta'}[L(x_i,x_j|\theta')-C(\theta')]\leq L(x_H,x_L|\theta)-C(\theta)=\Pi(x_H,x_L)$. $\{\theta,(x_H,x_L)\}$ is thus a Nash equilibrium strategy iff the payment scheme satisfies constraints (2-9) and (2-10). It follows from the discussion above that formulations [I-P] and [I-P'] are equivalent.

Proof of Lemma 2.2.

1)straightforward.

2) i) Since the payment scheme satisfies constraint (2-10), we have $\Pi(x_H,x_L)=L(x_H,x_L|\theta)-C(\theta)$. $\Pi(x_H,x_L)\geq\Pi(x_H,x_H)$ implies

$$\theta\pi_L[V_A(w(x_L,r_L))-V_A(w(x_H,r_L))]\geq(1-\theta)\pi_H[V_A(w(x_H,r_H))-V_A(w(x_L,r_H))]+C(\theta) \quad (A.1)$$

and $\Pi(x_H,x_L)\geq\Pi(x_L,x_L)$ implies

$$\theta\pi_H[V_A(w(x_H,r_H))-V_A(w(x_L,r_H))]\geq(1-\theta)\pi_L[V_A(w(x_L,r_L))-V_A(w(x_H,r_L))]+C(\theta) \quad (A.2)$$

(A.1)×(1-θ(e))+(A.2)×θ(e) followed by simple rearrangement provides

$$\pi_H[V_A(w(x_Hr_H))-V_A(w(x_L,r_H))](2\theta-1)\geq C(\theta)$$

$\theta>1/2$ implies $w(x_Hr_H)>(w_L,r_H)$. Similar derivation provides $w(x_L,x_L)>w(x_H,x_L)$.

2) ii) For any $\theta'\in[1/2,1)$, part i) implies $L(x_L,x_H|\theta')$ is strictly decreasing in θ', and the result follows.

Proof of Lemma 2.3.

I))Being risk neutral, the investor will invest W_0 in the risky asset if $E[r|x,\theta] \geq R$ and zero amount otherwise. Simple calculation provides that $E[r|x_H,\theta]>R$ and $E[r|x_H,\theta]<R$ for any $\theta>1/2$. As will be shown, the equilibrium accuracy $\theta_{RP}>1/2$. Thus $\lambda(x_L)=0$ and $\lambda(x_H)=W_0$.

Following Grossman and Hart (1983), the solution can be derived by first solving for the optimal compensation scheme implementing a given level of accuracy θ, and then optimize over θ in the second step. The solution to the first step will easily follow once we establish the following two facts.

Fact 1: Let f(x) be a function with $f'(x)>0$ and $f''(x)<0$. If $af(x_1)+(1-a)f(x_2)=af(y_1)+(1-a)f(y_2)$ where $x_1>y_1>y_2>x_2$ and $a\in(0,1)$, then $ax_1+(1-a)x_2>ay_1+(1-a)y_2$.

To show this, let $b\in(0,1)$ be such that $ax_1+(1-a)x_2=by_1+(1-b)y_2$. The existence of b is obvious. Define two random variables X,Z, with density functions $g_X(X=x_1)=a, g_X(X=x_2)=1-a$ and $g_Z(Z=y_1)=b$, $g_Z(Z=y_2)=1-b$. X and Z differ by a single mean preserving spread (MPS). It follows that $af(x_1)+(1-a)f(x_2)<bf(y_1)+(1-b)f(y_2)$ (Rothschild and Stiglitz (1970)). Thus $b>a$ and $ax_1+(1-a)x_2=by_1+(1-b)y_2>ay_1+(1-a)y_2$.

Fact 2: $\lim_{\theta\to+1}C'(\theta)=+\infty$

This is shown as follows: if $\lim_{e\to+\infty}C(e)$ is finite , then $\mu=\sup_e C(e)$ is finite. Thus for $\forall\epsilon>0$ $\exists e$ s.t $C(e)>\mu-\epsilon$. Since C(e) is strictly convex, thus $C(e_1)-C(e) >(e_1-e)C'(e)$, for $\forall e_1>e$. Choose e_1 s.t $e_1-e=\epsilon/C'(e)$. Then $C(e_1)>C(e)+\epsilon>\mu$ which forms a contradiction. Thus $\lim_{e\to+\infty}C(e)=\lim_{\theta\to1}C(\theta)=\infty$. Use of strict convexity $\frac{C(\theta)}{C'(\theta)}<\theta-\frac{1}{2}$ implies

$\lim_{\theta \to 1} C'(\theta) = +\infty$.

The optimization problem corresponding to the first step is the following.

$$Min_{w(\cdot,\cdot)} E_{x,r}[w(x,r)|\theta]$$

$$\text{s.t} \quad \frac{d}{d\theta} L(x_H, x_L|\theta) = C'(\theta)$$

$$w(x,r) \geq 0 \qquad x \in \{x_L, x_H\} \qquad r \in \{r_L, r_H\}$$

We have replaced constraint (2-16) in the reduced problem by its first order condition and have dropped the individual rationality constraint. The validity follows from the concavity of $L(x_H,x_L|\theta)$-$C(\theta)$ and that the limited liability constraint guarantees individual rationality be satisfied. Since

$$E_r[w(x_H,r)|\theta] = \theta[\pi_H w(x_H,r_H) + \pi_L w(x_L,r_L)] + (1-\theta)[\pi_H w(x_L,r_H) + \pi_L w(x_H,r_L)]$$

and $\frac{d}{d\theta} L(x_H,x_L|\theta) = [\pi_H V_A(w(x_H,r_H)) + \pi_L V_A(w(x_L,r_L))] - [\pi_H V_A(w(x_L,r_H)) + \pi_L V_A(w(x_H,r_L))$

It follows from fact 1, for given $\pi_H V_A(w(x_H,r_H)) + \pi_L V_A(w(x_L,r_L))$, $\pi_H w(x_H, r_H) + \pi_L w(x_L,r_L)$ is minimized at $w(x_H,r_H)) = w(x_L,r_L)$. Similarly, $\pi_H w(x_L,r_H) + \pi_L w(x_H,r_L)$ is minimized at $w(x_L,r_H) = w(x_H,r_L)$ for a given $\pi_H V_A(w(x_L,r_H)) + \pi_L V_A(w(x_H,r_L))$. Thus, at optimal,

$$E_r[w(x,r)|\theta] = \theta w(x_H,r_H) + (1-\theta) w(x_H,r_L)$$

$$\frac{d}{d} L(x_H,x_L|\theta) = V_A(w(x_H,r_H)) - V_A(w(x_H,r_L)) = C'(\theta)$$

The limited liability constraint for $w(x_H,r_L)$ must bind. If otherwise, a simultaneous decrease in $w(x_H,r_H)$ and $w(x_H,r_L)$ while leaving $V_A(w(x_H,r_H)) - V_A(w(x_H,r_L))$ unchanged will result in a decrease in $E_r[w(x,r)|\theta]$. We conclude that the optimal compensation scheme

implementing a given $\theta \geq 1/2$ must be such that $w(x_H, r_H)) = w(x_L, r_L) = h(C'(\theta))$ and $w(x_L, r_H) = w(x_H, r_L) = 0$[1].

ii) In the second step, the investor optimize over θ using the payment scheme derived in step one, i.e $Max_{\theta \in [1/2,1)}[\theta\beta - \theta h(C'(\theta))]$. Fact 2 implies, in optimizing over θ, we can restrict the range of θ to $[1/2, M]$ $M<1$. The existence of θ_{RP} follows from the continuity of $[\theta\beta - \theta h(C'(\theta))]$ in θ. The first order derivative of $[\theta\beta - \theta h(C'(\theta))]$ with respect to θ is strictly positive at $\theta = 1/2$. This implies $\theta_{RP} > 1/2$. To show $\theta_{RP} < \theta_{FB}$, notice $C(\theta = 1/2) = 0$ and $C''(\theta) > 0$ imply $C(\theta) < (\theta - 1/2)C'(\theta) < C'(\theta)$. $C'(\theta = 1/2) = 0$ and $C'''(\theta) \geq 0$ imply $C'(\theta) < (\theta - 1/2)C''(\theta) \leq \theta C''(\theta)$. Thus

$$\frac{d}{d\theta}[\theta h(C'(\theta))] = h(C'(\theta)) + \theta C''(\theta) h'(C'(\theta)) > \frac{d}{d\theta}[h(C(\theta))]$$

Recall, θ_{RP} and θ_{FB} satisfy

$$\frac{d}{d\theta}[h(C(\theta))]\Big|_{\theta=\theta_{FB}} = \beta = \frac{d}{d\theta}[\theta h(C'(\theta))]\Big|_{\theta=\theta_{RP}} > \frac{d}{d\theta}[h(C(\theta))]\Big|_{\theta=\theta_{RP}}$$

The monotonicity of $h(C(\theta))$ in θ implies $\theta_{RP} < \theta_{FB}$

Proof of Lemma 2.4.2.

From lemma 2.4.1, the optimization problem can be simplified as follows.

$$Min_{w(\cdot,\cdot)}\{\theta[\pi_H w(x_H, r_H) + \pi_L h(\frac{C(\theta) - \pi_H V_A(w(x_H, r_H))}{\pi_L}]\}$$

$$\pi_H V_A(w(x_H, r_H)) \in [(1-\theta)C'(\theta) + C(\theta), \theta C'(\theta) - C(\theta)]$$

[1] At $\theta = 1/2$, $w(x,r) = 0 \ \forall \ x,r$. In this case, no contracting occurs.

$$\pi_L V_A(w(x_L,r_L))=C'(\theta)-\pi_H V_A(w(x_H,r_H))$$

It can be easily verified that the objective function in the simplified problem is strictly convex. The strict convexity of $C(\theta)$ implies the closed interval constraining $\pi_H V_A(w(x_H,r_H))$ is nonempty. The continuity of the objective function in $w(x_H,r_H)$ and the continuity of h implies the solution to the reduced problem exists. Straightforward calculation indicates the first order derivative evaluated at the right end of the interval is strictly negative if and only if $\theta<\theta^0_H$. Strict convexity of the objective function implies that the constraints are not binding and the first order condition is sufficient if $\theta\geq\theta^0_H$. The first order condition provides $w(x_H,r_H))=w(x_L,r_L)=h(C'(\theta))$. Strict convexity also implies the objective function reaches minimum at $\pi_H V_A(w(x_H,r_H))=\theta C'(\theta)-C(\theta)$ when $\theta<\theta^0_H$.

Proof of Proposition 2.2.

If $\theta_{RP}\geq\theta^0_H$, then $\theta^0_H<1$. Lemma 2.4.2 implies that the objective function is $[\theta\beta-\theta h(C'(\theta))]$ for $\theta\geq\theta^0_H$ and is $\theta\{\beta-[\pi_H h(a_1(\theta)/\pi_H)+h(a_2(\theta)/\pi_L)\pi_L]\}$ for $\theta<\theta^0_H$. The strict convexity of h implies $\pi_H h(a_1(\theta)/\pi_H)+h(a_2(\theta)/\pi_L)\pi_L>h(C'(\theta))$. It follows

$$Max_{\theta\in[\frac{1}{2},\theta^0_H]}[\beta\theta-\theta(\pi_H h(\frac{a_1(\theta)}{\pi_H})+\pi_L h(\frac{a_2(\theta)}{\pi_L})]<Max_{\theta\in[\frac{1}{2},1)}[\beta\theta-\theta h(C'(\theta))]$$

$$= Max_{\theta\in[\theta^0_H,1)}[\beta\theta-\theta h(C'(\theta))]$$

Thus $\theta_{SB}=\theta_{RP}$.

If $\theta^0_H>\theta_{RP}$, lemma 2.4.2 indicates that the objective function is $[\theta\beta-\theta h(C'(\theta))]$ if $\theta\geq\theta^0_H$, and is $[\beta\theta-\theta(h(\frac{a_1(\theta)}{\pi_H})\pi_H+\pi_L h(\frac{a_2(\theta)}{\pi_L}))]$ for $\theta<\theta^0_H$. $[\theta\beta-\theta h(C'(\theta))]$ is strictly

concave in θ. $\theta_{RP}<\theta^0_H$ implies $\frac{d}{d\theta}[\theta\beta-\theta h(C'(\theta))]|_{\theta=\theta^0_H}<0$. On the other hand,

$$[\beta\theta-\theta(h(\frac{a_1(\theta)}{\pi_H})\pi_H+\pi_L h(\frac{a_2(\theta)}{\pi_L}))]|_{\theta=\theta^0_H} = [\theta\beta-\theta h(C'(\theta))]|_{\theta=\theta^0_H} > [\theta\beta-\theta h(C'(\theta))]$$

for any $\theta>\theta^0_H$. Thus, θ_{SB} must be the solution to the problem

$$Max_{\theta\in[\frac{1}{2},\theta^0_H]} \quad [\ \theta-\theta(h(\frac{a_1(\theta)}{\pi_H})\pi_H+\pi_L h(\frac{a_2(\theta)}{\pi_L}))]^2.$$

Furthermore, the first order derivatives of the objective function evaluated at $\theta=1/2$ and $\theta=\theta^0_H$ are strictly positive and negative respectively. Thus $1/2<\theta_{SB}<\theta^0_H$.

Proof of Proposition 2.3.

If θ^0_H is smaller than 1, then both of optimization problems in proposition 2.2 involves optimizing continuous functions over compact sets. The solution spaces of the two problems are thus compact subsets of real line, and a maximum exists in either case.

If either θ^0_H is 1, then fact 2 in the proof of lemma 2.3 can be used to restrict the range of θ to a compact set in the optimization problem proposition 2.2. The previous argument can again be applied to yield a compact solution space which contains a maximum.

Proof of Proposition 2.5.

(i)This follows directly from part i) of proposition 2.2.

(ii)If $\theta_{RP}<\theta^0_H$ and $\theta_{SB}\leq\pi_H$, the necessary condition for θ_{SB} to be the optimal accuracy level

---

[2]If $\theta_H{}^0=1$, the range of the θ is open on the right. In this case, fact 2 in the proof of lemma 2.3 can be used to show the existence of the solution.

is

$$\beta=\frac{d}{d\theta}\{\theta[\pi_H h(\frac{a_1(\theta)}{\pi_H})+\pi_L h(\frac{a_2(\theta)}{\pi_L})]\}|_{\theta=\theta_{SB}} .$$

The strict convexity of h implies

$$[\pi_H h(\frac{a_1(\theta)}{\pi_H})+\pi_L h(\frac{a_2(\theta)}{\pi_L})]|_{\theta=\theta_{SB}} > h(C'(\theta)) \ |_{\theta=\theta_{SB}}$$

The condition $h'''(\cdot)\geq 0$ implies

$$\frac{d}{d\theta}[\pi_H h(\frac{a_1(\theta)}{\pi_H})+\pi_L h(\frac{a_2(\theta)}{\pi_L})]|_{\theta=\theta_{SB}} \geq C''(\theta)h'(\theta\frac{a_1(\theta)}{\pi_H}+(1-\theta)\frac{a_2(\theta)}{\pi_L})|_{\theta=\theta_{SB}}$$

$$\geq C''(\theta)h'(C'(\theta) \ |_{\theta=\theta_{SB}} .$$

The second inequality follows from the assumption $\theta_{SB}\leq\pi_H$. Thus

$$\frac{d}{d\theta}(\theta h(C'(\theta))|_{\theta=\theta_{RP}}=\beta=\frac{d}{d\theta}\{\theta[\pi_H h(\frac{a_1(\theta)}{\pi_H})+\pi_L h(\frac{a_2(\theta)}{\pi_L})]\}|_{\theta=\theta_{SB}}>\frac{d}{d\theta}(\theta h(C'(\theta))|_{\theta=\theta_{SB}}$$

Since $\frac{d}{d\theta}(\theta h(C'(\theta))$ is strictly increasing in θ, we conclude $\theta_{SB}<\theta_{RP}$.

Proof of Proposition 2.6.

I)Since the expected total payoff from the investment is increasing in θ and that the agent does not earn any rents under the first best solution, Lemma 4 implies $P_{FB}>P_{RP}$.

To compare P_{SB} with P_{RP}, we consider the following three cases.

Case 1. $\theta_{RP}\geq\theta^0_H$, Proposition 2.2 implies $P_{SB}=P_{RP}$,

Case 2. $\theta^0_H>\theta_{RP}$. The profit for the investor is

$$\beta\theta-\theta(\pi_H h(\frac{a_1(\theta)}{\pi_H})+\pi_L h(\frac{a_2(\theta)}{\pi_L}))]|_{\theta=\theta_{SB}} < [\beta\theta-\theta h(a_1(\theta)+a_2(\theta))]|_{\theta=\theta_{SB}}$$

$$= [\beta\theta-\theta h(C'(\theta))]|_{\theta=\theta_{SB}} .$$

When the signal is observable, the investor derives a profit $[\beta\theta-\theta h(C'(\theta))]|_{\theta=\theta_{RP}}$. Since θ_{RP} solves $Max_\theta[\beta\theta-\theta h(C'(\theta))]$, we have

$$[\beta\theta-\theta h(C'(\theta))]|_{\theta=\theta_{RP}} \geq [\beta\theta-\theta h(C'(\theta))]|_{\theta=\theta_{SB}} .$$

The investor derives a strictly smaller profit in the second best case than that in the reduced problem, i.e $P_{SB}<P_{RP}$.

ii)For fixed β, define π_H^c by $[\theta-\frac{C(\theta)}{C'(\theta)}]|_{\theta=\theta_{RP}}=\pi_H^c$, where θ_{RP} is the optimal accuracy level in the reduced problem. Following Proposition 2.2, the investor's profit is a constant P_s for $\pi_H\leq\pi_H^c$, i.e $\theta^0{}_H\leq\theta_{RP}$. For $\pi_H>\pi_H^c$, the optimal payment scheme is given by Proposition 2.2. Let $s(\theta,\pi_H)= \beta\theta-\theta[\pi_H h(\frac{a_1(\theta)}{\pi_H})+\pi_L h(\frac{a_2(\theta)}{\pi_L})]$, then the investor's profit is the following expression

$$P_a(\pi_H)=Max_{\theta\in[\frac{1}{2},\theta_H^0(\pi_H)]}s(\theta,\pi_H) ,$$

where $\theta_H{}^0(\pi_H)$ is defined as in Definition 2.1.

Consider two priors $\pi_H{}^1$ and $\pi_H{}^2$, such that $\pi_H{}^c<\pi_H{}^1<\pi_H{}^2$. By definition, $\theta_H{}^0(\pi_H{}^1)<\theta_H{}^0(\pi_H{}^2)$. Further it is easily seen that $s(\theta,\pi_H{}^1)\geq s(\theta,\pi_H{}^2)$, $\forall\theta\leq\theta_H{}^0(\pi_H{}^1)$. Thus, if $P_a(\pi_H{}^2)>P_a(\pi_H{}^1)$, then $s(\theta,\pi_H{}^2)$ must be maximized at some point θ_2 such that $\theta_H{}^0(\pi_H{}^1)<\theta_2\leq\theta_H{}^0(\pi_H{}^2)$. Since $\theta_{RP}<\theta_H{}^0(\pi_H{}^2)$, it follows that $\frac{d}{d\theta}s(\theta,\pi_H^2)|_{\theta=\theta_H(\pi_H^2)}<0$. Thus

$\theta_2<\theta_H^0(\pi_H^2)$. Now, define prior π_H^3 by $\left[0-\frac{C(\theta)}{C'(\theta)}\right]|_{\theta=\theta_2}=\pi_H^3$. It follows that $\pi_H^1<\pi_H^2<\pi_H^3$. Since $s(\theta_2,\pi_H^2)=P_a(\pi_H^2)$ and $P_a(\pi_H^3)>s(\theta_2,\pi_H^3)$, we have $P_a(\pi_H^3)>P_a(\pi_H^2)>P_a(\pi_H^1)$. We can continue to apply the procedure and construct the sequence $\{\pi_H^n, n=2,\cdot,+\infty\}$ with $\pi_H^1\le\pi_H^{n+1}<\pi_H^n$. Thus, the sequence must converge to $\pi_H'\ge\pi_H^1$. From the procedure of constructing the sequence, $P_a(\pi_H^n)=s(\theta_H^0(\pi_H^{n+1}),\pi_H^n)$. From Definition 2.1, $\theta_H^0(\pi_H)$ is continuous in π_H. The continuity of the function $s(\theta,\pi_H)$ implies that

$$\lim_{n\to\infty}P_a(\pi_H^n)=s(\theta_H^0(\pi_H'),\pi_H').$$

Since $\pi_H'\ge\pi_H^1$, it follows that $\theta_H^0(\pi_H')>\theta_{RP}$. It follows from this inequality that $P_a(\pi_H')>s(\theta_H^0(\pi_H'),\pi_H')>P_a(\pi_H^1)$. Let $s(\theta,\pi')$ be maximized at θ_0, i.e $P_a(\pi_H')=s(\theta_0,\pi_H')$. Then, $P_a(\pi_H')\ge P_a(\pi_H(n))\ge s(\theta_0,\pi_H(n))$ implies $\lim_{n\to\infty}P_a(\pi_H(n))=P_a(\pi_H')$. This then forms a contradiction. Thus, we conclude that $P_a(\pi_H^2)\le P_a(\pi_H^1)$.

APPENDIX B

PROOFS OF THE MAIN RESULTS IN CHAPTER 3

Proof of Proposition 3.1

The assumptions on $P(q,e,s)$ and on $C(q)$ and $D(e)$ suffice to insure $U(q,e,s)$ is strictly concave in q and $\Pi(q,e,s)$ is strictly concave in e. If we further require that $q,e \leq \mu < \infty$ then by Theorem 3.1 of Friedman (1990), a pure strategy Nash equilibrium exists. The conditions on *P(q,e,s)*, *C(q)* and *D(e)* further insure that the Nash equilibrium is interior (with $e,q > 0$) and that it is characterized by the first order conditions (3-1) and (3-2) in the text. Finally, uniqueness of equilibrium follows by verifying that the reaction function of the party and of the enforcer are continuous and have slopes of opposite signs indicating a unique equilibrium at the single point of intersection.

Proof of Proposition 3.2

Totally differentiating eqs. (3.1) and (3.2) in the text with respect to s yields the following:

$$\begin{pmatrix} -P_{qq}-C_{qq} & -P_{qe} \\ P_{qe} & P_{ee}-D_{ee} \end{pmatrix} \begin{pmatrix} dq/ds \\ de/ds \end{pmatrix} = \begin{pmatrix} P_{qs} \\ -P_{es} \end{pmatrix} \qquad \text{(B.1)}$$

Cramer's rule applied to (C.1) implies:

$$\frac{dq}{ds} = \frac{P_{qs}\,(P_{ee} - D_{ee}) \;-\; P_{qe}P_{es}}{\Delta} \;>\; 0 \qquad \text{(B.2)}$$

$$\frac{de}{ds} = \frac{(P_{qq} + C_{qq})P_{es} \;-\; P_{eq}P_{qs}}{\Delta} \;\begin{matrix}>\\<\end{matrix}\; 0 \;\; as \;\; P_{qe} \;\begin{matrix}>\\<\end{matrix}\; 0 \qquad \text{(B.3)}$$

where $\Delta = -(P_{qq} + C_{qq})(P_{ee} - D_{ee}) + P_{qe}^2 > 0$. The sign of dq/ds follows immediately from our assumptions about P and D. To verify the sign of de/ds, rewrite (B.3) so that

$$\frac{de}{ds} = \left\{\frac{P_{se}}{P_{qe}} - \frac{P_{qs}}{(P_{qq} + C_{qq})}\right\} \frac{(P_{qq} + C_{qq})\,P_{qe}}{\Delta} \qquad \text{(B.4)}$$

$$= \left\{\left.\frac{dq}{ds}\right|_{de=0} - \left.\frac{dq(e(s),s)}{ds}\right|_{de(s)=0}\right\} \frac{(P_{qq} + C_{qq})\,P_{qe}}{\Delta}$$

$$\begin{matrix}>\\<\end{matrix}\; 0 \;\; as \;\; P_{qe} \;\begin{matrix}>\\<\end{matrix}\; 0 \iff P_{se} \;\begin{matrix}<\\>\end{matrix}\; 0$$

by Assumption 3.1.

Proof of Proposition 3.3

The optimal standard $\tilde{s}$ satisfies

$$\frac{dV}{d\tilde{s}} = (B - C_q)\frac{dq}{ds} - D_e\frac{de}{ds}\;0 \qquad \text{(B.5)}$$

Solving for $(B - C_q)$ from (B.5) yields

$$B - C_q = \frac{D_e \dfrac{de}{ds}}{\dfrac{dq}{ds}} \tag{B.6}$$

It follows from (B.6) and Proposition 3.2 that

$$B - C_q \gtrless 0 \text{ as } \frac{de}{ds} \gtrless 0 \iff P_{se} \gtrless 0 \tag{B.7}$$

Finally, since $C_{qq} < 0$, and q is increasing in *s*, it follows from (B.7) that

$$\tilde{s} \gtrless s^* \text{ as } P_{se} \gtrless 0 \tag{B.8}$$

Proof of Proposition 3.4

First we provide necessary and sufficient conditions for satisfying the conditions (i) and (ii) of [GP-A] in the text. Applying routine arguments (see Guesnerie and Laffont (1984)) one can readily show that the schedules $\{T(\theta),\ s(\theta)\}$ are differentiable almost everywhere, and that the effort level induced, $e(s(\theta'),\ \theta)$ must be non-increasing in θ' where $\theta' = \theta$. Further,

$$\Pi'(\theta) = \Pi_1(\theta' \mid \theta)\Big|_{\theta'=\theta} \frac{d\theta'}{d\theta} + \Pi_2(\theta'|\theta)\Big|_{\theta'=\theta} \tag{B.9}$$

$$= -D_0\,(e(s(\theta),\theta),\theta)$$

where the second line of (B.9) follows from the Envelope Theorem. Since $\Pi(\theta)$ is decreasing, part (i) is insured provided

$$\Pi(\bar{\theta}) = 0 \tag{B.10}$$

Combining (B.9) - (B.10), parts (i) and (ii) of [GP-A] are satisfied provided,

$$\Pi(\theta) = \int_{\theta}^{\bar{\theta}} - D_{\theta}\,(e(s(\tilde{\theta}),\tilde{\theta}),\tilde{\theta})\,d\tilde{\theta} \tag{B.11}$$

Substituting for $\Pi(\theta)$ from (C.11) into [GP-A], integrating by parts and rearranging terms yields

$$\max_{s(\theta)} E_{\theta}\,\hat{V}(s(\theta),\theta) = \max_{s(\theta)} E_{\theta}\left\{B_q(\cdot) - C(q(\cdot)) - D(e(\cdot),\theta) - (1-\lambda)\,D_{\theta}\,(e(\cdot),\theta))\ \frac{F(\theta)}{f(\theta)}\right\}$$

where we have deleted the arguments of $q(\cdot)$ *and* $e(\cdot)$ for notational convenience. Rewriting $\hat{V}\,(s(\theta),\theta)$ in terms of $V(s(\theta),\theta) \equiv B_q(\cdot) - C\,(q(\cdot)) - D\,(e(\cdot),\theta)$ and recognizing that D_{θ} is implicitly a function of $s(\theta)$ and θ we have

$$\hat{V}\,(s(\theta),\theta) = V\,(s(\theta),\theta) - (1-\lambda)\,D_{\theta}\,(s(\theta),\theta)\ \frac{F(\theta)}{f(\theta)} \tag{B.12}$$

Assuming a separating solution to [GP-A], Let

$$\hat{s}\,(\theta) = \underset{s(\theta)}{argmax}\ \hat{V}\,(s(\theta),\theta)$$

$$\hat{s}\,(\theta) = \underset{s(\theta)}{argmax}\ V\,(s(\theta),\theta)$$

Then, employing standard revealed preference arguments for all $\theta \in [\underline{\theta},\bar{\theta}]$

$$\hat{V}\,(\hat{s}(\theta),\theta) \geq \hat{V}\,(\bar{s}(\theta),\theta) \tag{B.13}$$

$$V\,(\bar{s}(\theta),\theta) \geq V\,(\hat{s}(\theta),\theta) \tag{B.14}$$

with strict inequality for $\theta > \underline{\theta}$. Adding (B.13) and (B.14) and simplifying yields

$$(1 - \lambda)[D_\theta(\tilde{s}(\theta),\theta) - D_\theta(\hat{s}(\theta),\theta)] \geq 0 \tag{B.15}$$

This implies, since $1 - \lambda > 0$ that

$$\tilde{s}(\theta) \gtrless \hat{s}(\theta) \text{ as } \frac{d}{ds(\theta)} D_\theta\,(s(\theta),\theta) \gtrless 0 \tag{B.16}$$

But

$$\frac{d}{ds(\theta)} D_\theta(s(\theta),\theta) = D_{\theta e} \frac{de}{ds} \tag{B.17}$$

$$\text{sgn } \frac{de}{ds} \gtrless 0 \text{ as } P_{sq} \gtrless 0$$

where the second line of (B.17) follows from Proposition 3.2. Collecting (B.16) and (B.17) we have

$$\hat{s}(\theta) \leq \tilde{s}(\theta) < s^* \quad for \quad P_{sq} > 0 \tag{B.18}$$

$$\hat{s}(\theta) \geq \tilde{s}(\theta) > s^* \quad for \quad P_{sq} < 0 \tag{B.19}$$

with strict inequality for $\theta > \underline{\theta}$, thus proving Proposition 3.4.

Proof of Proposition 3.5

The solution to [GP-P] as posed in the text is characterized by the first order conditions

$$E_{\mu \leq \hat{\mu}} [B - C_q(q(\mu,s),\mu)] \frac{dq(\cdot)}{ds} - D'(F(\hat{\mu})e(s)) \frac{de}{ds} = 0 \tag{B.20}$$

$$\hat{V}(\hat{\mu}) \equiv B_{q(\hat{\mu},s)} - C(q(\hat{\mu},s),\hat{\mu}) - D'(F(\hat{\mu})\,e(s))e(s) \geq 0 \quad (= \text{if } \hat{\mu} < \bar{\mu}) \quad \text{(B.21)}$$

where (B.20) and (B.21) correspond respectively to the maximization of [GP-P] with respect to s and $\hat{\mu}$. The Nash equilibrium care and enforcement levels, $q(\mu,s)$ *and* $e(s)$ are characterized by

$$-P_q(q(\mu,s),e(s),s) - C_q(q(\mu,s),\mu) = 0;\ \mu \leq \hat{\mu} \quad \text{(B.22)}$$

$$E_{\mu \leq \hat{\mu}}\, P_e(q(\mu,s),e(s),s) - D'(F(\hat{\mu})\,e(s))\,F(\hat{\mu}) = 0 \quad \text{(B.23)}$$

First we prove parts (iii) and (iv) of the Proposition. Differentiating (B.22) and (B.23) totally w.r.t. s yields:

$$-P_{qq}\left(\frac{dq}{ds}\right) - P_{qe}\left(\frac{de}{ds}\right) - P_{qs} - C_{qq}\left(\frac{dq}{ds}\right) = 0;\quad \mu \leq \hat{\mu} \quad \text{(B.24)}$$

$$E_{\mu \leq \hat{\mu}}\left\{P_{eq}\left(\frac{dq}{ds}\right) + P_{ee}\left(\frac{de}{ds}\right) + P_{es}\right\} - D''F(\hat{\mu})\left(\frac{de}{ds}\right) = 0 \quad \text{(B.25)}$$

Combining (B.24) and (B.25) one obtains

$$\frac{de}{ds} = \frac{A}{B} \quad \text{(B.26)}$$

where

$$A = -E_{\mu \leq \hat{\mu}}\left\{\frac{-P_{qs}}{P_{qq} + C_{qq}} + \frac{P_{es}}{P_{qe}}\right\} P_{eq} \quad \text{(B.27)}$$

$$B = E_{\mu \le \hat{\mu}} \left\{ \frac{-P_{qe}^2}{P_{qq} + C_{qq}} + P_{ee} \right\} - D'F(\hat{\mu}) < 0 \qquad \text{(B.28)}$$

It follows from Assumption 3.1 and (B.26) - (B.28) that

$$\frac{de}{ds} \gtrless 0 \text{ as } P_{qe} \gtrless 0 \iff P_{qs} \gtrless 0 \qquad \text{(B.29)}$$

In addition (B.24) and (B.25) also imply that

$$\frac{dq(\mu,s)}{ds} = -\frac{P_{qe}\left(\frac{de}{ds}\right) + P_{qs}}{(P_{qq} + C_{qq})} > 0 \qquad \text{(B.30)}$$

where the inequality follows from (B.29). Substituting (B.29) and (B.30) into the first order condition for s, (B.23) allows one to verify parts (iii) and (iv) of Proposition 3.5.

To verify part (i), notice that for B sufficiently large $V(\hat{\mu})$ is strictly positive for all $\hat{\mu}$, as both terms $C(q(\hat{\mu},s),\hat{\mu})$ and $D'(\cdot)\,e(s)$ in (B.21) are bounded above since $q(\mu,s)$ and $e(s)$ are bounded, while the term $B\,q(\hat{\mu},s)$ is arbitrarily large. Hence, $\hat{\mu} = \bar{\mu}$ and no party types are exempted when the marginal benefits of care are sufficiently large.

To verify part (ii), notice that for a type which decides to exempt himself

$$-A \ge \min_{q} \left(-p(q,e(s),s) - C(q,\mu)\right) > -F \qquad \text{(B.31)}$$

or A < F. This completes the proof of Proposition 3.5.

Proof of Proposition 3.6

Given s, the first order condition for Y in the solution to the enforcer's problem, [EP] stated in the text is

$$(P(q(\mu',s),e(s),s) - D'(F(\mu')e(s))e(s) - Y)f(\mu')\left(\frac{d\mu'}{dY}\right) + (1 - F(\mu')) \geq 0$$

$$(= \text{if } \mu' < \bar{\mu}) \tag{B.32}$$

Notice that when $\mu' = \bar{\mu}$

$$Y = P(q(\bar{\mu},s),e(s),s) + C(q(\bar{\mu},s),\bar{\mu}) \tag{B.33}$$

Substituting for this value of Y in (B.32) reveals that $\mu' = \bar{\mu}$ can not be a solution to [EP]. Therefore, $\mu' < \bar{\mu}$.

Rearranging (B.32) and noting that it holds with equality yields the expression appearing in Proposition 3.6. Finally, the result that Y < F follows from noting that

$$-Y \geq \min_{q} (-p(q,e(s),s) - C(q,\mu)) > -F \tag{B.34}$$

for all $\mu \leq \bar{\mu}$.

Proof of Proposition 3.7

According to eq (3-3) in the text

$$\frac{dV}{dF} \begin{matrix} > \\ = \\ < \end{matrix} 0 \quad as \quad \frac{de/ds}{dq/ds} \begin{matrix} > \\ = \\ < \end{matrix} \frac{de/dF}{dq/dF}$$

The expressions for $\frac{de/ds}{dq/ds}$ *and* $\frac{de/dF}{dq/dF}$ are given by

$$\frac{de/d\alpha}{dq/d\alpha} = \frac{(\lambda_\alpha + P_{eq})\,\Delta_q\,(P_{qe}^2 - P_{ee}\,\Delta q)}{(P_{qe}^2 - P_{ee}\,\Delta_q)^2 + (P_{qe}^2 - P_{ee}\,\Delta_q)(\lambda_\alpha + P_{eq})} \;;\; \alpha = S,F \tag{B.35}$$

where

$$\lambda_F = -\frac{P_e}{P_q}\,\Delta_q > 0 \tag{B.36}$$

$$\lambda_s = -\frac{P_{es}}{P_{eq}}\,\Delta_q > 0 \tag{B.37}$$

$$\Delta_q = P_{qq} - C_{qq} > 0 \tag{B.38}$$

It is easy to demonstrate that the RHS of (B.35) is increasing in λ_α so that

$$\frac{de/ds}{dq/ds} \gtreqless \frac{de/dF}{dq/dF} \iff \lambda_S \gtreqless \lambda_F$$

$$\iff -\frac{P_{es}}{P_{eq}} \gtreqless -\frac{P_e}{P_q} \tag{B.39}$$

$$\iff \frac{d}{ds}\left(-\frac{P_e}{P_q}\right) \gtreqless 0$$

where one can easily verify the last equivalence in (B.39).

APPENDIX C

PROOFS OF THE MAIN RESULTS IN CHAPTER 4

Proof of Lemma 4.1.

Assuming that the bank is uninformed, consider first that the firm makes offers. Let the equilibrium offer be $G(\theta_i)$, i=H,L. $G(\theta)$ is the offer made by the firm in state θ. Consider first separating equilibriums with $G(\theta_H) \neq G(\theta_L)$.

Case S.1. Project is liquidated at θ_H and continued at θ_L.

In this case, $G(\theta_H)$ represents a t=1 repayment to the bank and corresponds to a liquidation contract; $G(\theta_L)$ represents a t=2 repayment to the bank and is a continuation contract. In state θ_L, the firm's and the bank's individual rationality conditions requires that

$$R_f(G(\theta_L),t_2|\theta_L) \geq R_f(s_2,t_2|\theta_L),$$

$$R_b(G(\theta_L),t_2|\theta_L) \geq R_b(s_2,t_2|\theta_L). \qquad (C.1)$$

Since $R_f(x,t_2|\theta_L)$ is continuously decreasing in x, thus $G(\theta_H) \leq s_2$. $R_b(x,t_2|\theta_L)$ is continuously increasing in x, thus $G(\theta_L) \geq s_2$. Together, they imply $G(\theta_L)=s_2$. In state θ_H, the firm's individual rationality requires

$$L_b^{\ell} - G(\theta_H) \geq R_f(s_2,t_2|\theta_L) + C_L. \qquad (C.2)$$

In addition, incentive compatibility requires that the firm, in state θ_L, will not choose to liquidate the project by proposing the liquidation contract $G(\theta_H)$, i.e.

$$R_f(s_2,t_2|\theta_L) + C_L \geq L_b^{\ell} - G_1(\theta_H). \qquad (C.3)$$

Together, (C.2) and (C.3) imply

$$R_f(s_2,t_2|\theta_L)+C_L \geq L_b^t-G(\theta_H) \geq R_b(s_2,t_2|\theta_H)+C_H. \quad (C.4)$$

But, $R_f(s_2,t_2|\theta_H) \geq R_f(s_2,t_2|\theta_L)$ and $C(\theta_H)>C(\theta_L)$ imply that such a separating equilibrium is infeasible.

Case S.2 Project is liquidated in state θ_L and continued in state θ_H.

In this case, $G(\theta_H)$ represents a t=2 repayment to the bank and corresponds to a continuation contract; $G(\theta_L)$ represents a t=1 repayment to the bank and corresponds to a liquidation contract. Again, individual rationality conditions for the firm and the bank imply $G(\theta_H)=s_2$. In state θ_L, the liquidation contract must satisfy the following constraints,

$$L_b^t-G(\theta_L) \geq R_f(s_2,t_2|\theta_L)+C_L,$$

$$G(\theta_L) \geq R_b(s_2,t_2|\theta_L),$$

$$R_f(s_2,t_2|\theta_H)+C_H > L_b^t-G(\theta_L). \quad (C.5)$$

The first two constraints are the firm's and the bank's individual rationality conditions. The last constraint is the incentive compatibility constraint which ensures that the firm, in state θ_H, will prefer to continue than to liquidate the project. $G(\theta_L)$ satisfying these constraints exists if

$$L_b^t-[R_f(s_2,t_2|\theta_L)+C_L] \geq R_b(s_2,t_2|\theta_L). \quad (C.6)$$

If this inequality holds, separating equilibrium exits and

$$G(\theta_L)=\text{Max}\{R_b(s_2,t_2|\theta_L), L_b^t-[R_f(s_2,t_2|\theta_H)+C_H]\}. \quad (C.7)$$

Consider next pooling equilibrium with $G(\theta_H)=G(\theta_L)$.

Case P.1. Renegotiation leads to liquidation in both states.

In this case, both $G(\theta_H)$ and $G(\theta_L)$ are liquidation contracts and in equilibrium they

must be the same. We denote them by G_L. The individual rationality constraints for the firm and the bank are

$$L_b^{\ell}-G_L \geq R_f(s_2,t_2|\theta_H)+C_H,$$

$$G_L \geq E_\theta[R_b(s_2,t_2|\theta)]. \qquad (C.8)$$

G_L satisfying these constraints exits if

$$L_b^{\ell} \geq R_f(s_2,t_2|\theta_H)+C_H+E_\theta[R_b(s_2,t_2|\theta)]. \qquad (C.9)$$

If (C.9) holds, the manager proposes $G_L=E_\theta[R_b(s_2,t_2|\theta)]$.

Case P.2. Renegotiation leads to continuation in both interim states.

In this case, both $G(\theta_H)$ and $G(\theta_L)$ are continuation contracts and in equilibrium they must be the same. We denote them by G_c. Again, individual rationality constraints require $G_c=s_2$. The outcomes of this equilibrium is the same as that without interim renegotiation.

We establish the uniqueness in two steps.

Step 1: If (A.9) holds then the pooling equilibrium P.1 is the unique equilibrium.

In this case, both the separating equilibrium S.2 and the pooling equilibrium P.2 are feasible. We show, however, divinity criterion upsets these two equilibriums. Let G_1' be a liquidation contract satisfying

$$L_b^{\ell}-[R_f(s_2,t_2|\theta_H)+C_H]>G_1'>E_\theta[R_b(s_2,t_2|\theta)]. \qquad (C.10)$$

In the equilibrium S.2, G_1' is strictly preferred by the firm in both states. By divinity, bank's conjecture about the states, when faced with the offer G_1', is the same as the prior. Since $G_1'>E_\theta[R_b(s_2,t_2|\theta)]$, the bank will accept such an offer. This upsets the separating equilibrium. Same reasoning indicates that divinity criterion also upsets the pooling equilibrium P.2.

Step 2:If (C.6) holds and

$$L_b^L-[R_f(s_2,t_2|\theta_L)+C_L]\geq E_\theta[R_b(s_2,t_2|\theta)], \qquad (C.11)$$

then the separating equilibrium S.2 is the unique equilibrium.

In this case, pooling equilibrium P.1 is clearly infeasible. Suppose the equilibrium is the pooling equilibrium P.2. Let G_1' be a liquidation contract which satisfies

$$R_f(s_2,t_2|\theta_H)+C_H>L_b^L-G_1'\geq R_f(s_2,t_2|\theta_L)+C_L,$$

$$G_1'\geq R_b(s_2,t_2|\theta_L). \qquad (C.12)$$

If such a contract G_1' exists, the manager will deviate and offer the liquidation contract G_1' in state θ_L. By the divinity criterion and (C.12), the bank will accept such offer. This then upsets the proposed pooling equilibrium P.2. G_1' satisfying the two constraints exits if (C.6) holds.

When the bank makes an offer, it offers a menu $\{F(\theta_H),F(\theta_L)\}$, where $F(\theta_i)$ is the contract intended for the firm in state θ_i, i=H,L. From the above analysis, a menu, which induces liquidation in state θ_H and continuation in state θ_L, can not be incentive compatible. If the menu induces continuation in both state, then the bank's and the firm's individual rationality conditions imply that $F(\theta_H)=F(\theta_L)=s_2$. This is the same outcome as that without renegotiation. If the menu induces liquidation in both states, the firm's incentive compatibility condition requires that $F(\theta_H)=F(\theta_L)=F_L$. The firm's individual rationality condition requires $F_L\geq R_f(s_2,t_2|\theta_H)+C_H$. Thus the bank will choose to offer $F_L=R_f(s_2,t_2|\theta_H)+C_H$. Finally, the bank can offer a menu which induces liquidation in state θ_L and continuation in state θ_H. The firm's individual rationality condition requires that $F(\theta_H)=s_2$. The contract $F(\theta_L)$ must satisfy the incentive compatibility condition and the individual rationality condition for the firm,

$$F(\theta_L)\geq R_f(s_2,t_2|\theta_L)+C_L$$

$$R_f(s_2,t_2|\theta_H)+C_H \geq F(\theta_L). \tag{C.13}$$

$F(\theta_L)$ satisfying (C.13) exits if the condition (C.6) holds, in which case the bank will propose

$$F(\theta_L)=R_f(s_2,t_2|\theta_L)+C_L. \tag{C.14}$$

It is easily seen that when conditions (C.6) and (C.11) hold, the bank will propose the menu which induces liquidation in state θ_L and continuation in state θ_H.

Proof of Lemma 4.2.

Consider first the manager makes the offer. If the manager offers a separating contract, the bank can perfectly infer the realized state. Since $L>r(\theta_L)$, the bank will refuse the offer indicating the state θ_L and liquidate the firm. Thus the equilibrium must be pooling. By assumption 4.1, the firm can offer a continuation contract s_2^r, in both states, such that

$$E_\theta[s_2^r-\int_0^{t_2^r}F(r|\theta)]=L_b. \tag{C.15}$$

This contract satisfies the bank's individual rationality constraint and will be accepted by bank. The project is always continued.

Consider next the bank makes the offer. The uninformed bank offers a menu $\{G(\theta_H),G(\theta_L)\}$. Assuming first that the menu is separating, i.e. $G(\theta_H)\neq G(\theta_L)$. It is easily seen that a separating menu which induces continuation in state θ_L and liquidation in state θ_H can not be incentive compatible. Focusing on the opposite case, the incentive compatibility constraints are

$$C_H+\int_{G(\theta_H)}^{X}(r-G(\theta_H))dF(r|\theta_H)>L-G(\theta_L), \tag{C.16}$$

$$L-G(\theta_L)\geq C_L+\int_{G(\theta_H)}^{X}(r-G(\theta_H))dF(r|\theta_L). \qquad (C.17)$$

The first inequality ensures, in state θ_H, the firm will choose the contract $G(\theta_H)$ rather than $G(\theta_L)$. The second ensures, at θ_L, the firm chooses the contract $G(\theta_L)$. Combining the two constraints provides

$$C_H+\int_{G(\theta_H)}^{X}(r-G(\theta_H))dF(r|\theta_H)>L-G(\theta_L)\geq C_L+\int_{G(\theta_H)}^{X}(r-G(\theta_H))dF(r|\theta_L). \qquad (C.18)$$

By assumption 4.3, (C.18) can not be satisfied. It follows that the bank proposes a pooling menu, i.e.$G(\theta_H)=G(\theta_L)$. By assumption 4.1, the bank proposes a pooling contract demanding the entire t=2 cash flow, i.e. $s_2^b=X$.

Proof of Proposition 4.3

Since we only need to show existence, we prove the proposition by constructing such a repayment schedule. We require (s_1^*,s_2^*) satisfy the following conditions.

Condition 1: $$(1-\lambda)\int_0^X rdF(r|\theta_H)+\lambda L>s_1^*>(1-\lambda)E_\theta\int_0^X rdF(r|\theta)+\lambda L \quad . \qquad (C.19)$$

Condition 2: $$s_1^*=(D_m^i(\theta_H)+s_2^*)-\int_0^{D_m^i(\theta_H)+s_2^*}F(r|\theta_H)dr \qquad (C.20)$$

$$I_m^i(\theta_H)=D_m^i(\theta_H)-\int_0^{D_m^i(\theta_H)}F(r|\theta_H)dr \qquad (C.21)$$

where $D_m^i(\theta_H)>0$ and $s_1^*-I_m^i(\theta_H)<L$.

Condition 3: s_1^* is senior with seniority protected. s_2^* is junior and allows the firm to issue additional public debt up to $D_m^i(\theta_H)$.

From assumption 4.1, s_1^* and s_2^* satisfying the conditions exist. We claim that, in the

unique equilibrium, the firm raises $I_m^i(\theta_H)$ by issuing public debt with face value $D_m^i(\theta_H)$ and senior to $s_2^*$when the bank is informed of the favorable state; When the bank is uninformed, the project is continued through renegotiation. The project is liquidated when the bank is informed of the state θ_L. We prove this claim in two steps.

Step 1: The firm will not raise $I_m^i(\theta_H)$ from the market except when the bank is informed of the state θ_H.

Consider first the case when the bank is uninformed. Suppose the firm deviates and issues the senior public debt with face value $D_m^i(\theta_H)$, then

$$s_1^* - I_m^i(\theta_H) = s_2^* - \int_{D_m^i(\theta_H)}^{D_m^i(\theta_H)+s_2^*} F(r|\theta_H)dr > E_\theta[s_2^* - \int_{D_m^i(\theta_H)}^{D_m^i(\theta_H)+s_2^*} F(r|\theta)dr] \quad . \qquad (C.22)$$

The right hand side of (C.22) is the bank's expected profit from the promised t=2 repayment, s_2^*, if the project is continued. The left hand side is no greater than the bank's payoff if the project is liquidated after the repayment $I_m^i(\theta_H)$. Thus, when the bank is uninformed, it will not forgive $s_1^* - I_m^i(\theta_H)$. The firm must still negotiate with the bank for the additional financing. The firm is strictly better off acquiring the entire additional financing through negotiating with the bank. Suppose, the firm raises I_m^u such that $s_1^* - I_m^u$ is the same as the bank's expected payoff in continuation. Then, the bank will forgive $s_1^* - I_m^u$ and allow the project to continue. The firm's profit is

$$\int_0^X r dF(r|\theta) - (s_1^* - I_m^u) - I_m^u. \qquad (C.23)$$

By condition 1, one can easily see that the firm is strictly better off negotiating with the bank. Thus, when the bank is uninformed, the firm never raises capital from the market.

Consider next the case when the bank is informed of the state θ_L. If the firm raises $I_m^i(\theta_H)$, then

$$s_1^* - I_m^i(\theta_H) = s_2^* - \int_{D_m^i(\theta_H)}^{D_m^i(\theta_H)+s_2^*} F(r|\theta_H)dr > s_2^* - \int_{D_m^i(\theta_H)}^{D_m^i(\theta_H)+s_2^*} F(r|\theta_L)dr \ . \qquad (C.24)$$

Thus, the bank will not forgive s_1^*. Since $s_1^* > L > s_1^* - I_m^i(\theta_H)$, it is clear that the bank will liquidate the firm. Consequently, the firm will not raise $I_m^i(\theta_H)$ in this case. Thus, if the firm attempts to raise capital from the market, it will make an offer other than $I_m^i(\theta_H)$. It follows then the market will know the realization of θ_L. Since $s_1^* > L$, by assumption 0, the project must be liquidated when the bank is informed of the state θ_L.

Step 2. We show that the unique equilibrium strategy for the firm in state θ_H and with the bank informed is to raise senior debt $\{I_m^i(\theta_H), D_m^i(\theta_H)\}$ from the market.

If the firm raises $I_m^i(\theta_H)$ and repay the bank, the bank's payoff in liquidation is $s_1^* - I_m^i(\theta_H)$ and is the same as its payoff if the project is continued under the exiting contract. Thus the bank will forgive $s_1^* - I_m^i(\theta_H)$. The firm's profit in this case is

$$\int_0^X r dF(r|\theta_H) - [s_1^* - I_m^i(\theta_H)] - I_m^i(\theta_H) \quad . \qquad (C.25)$$

Alternatively, the firm can acquire financing by only negotiate with the bank and get

$$\lambda[\int_0^X r dF(r|\theta_H) - L] \ . \qquad (C.26)$$

From condition 1, the firm strictly prefers to raise $I_m^i(\theta_H)$ from the market.

REFERENCES

Aghion, P., and P. Bolton, (1992), " An Incomplete Contracts Approach to Financial Contracting," *Review of Economic Studies* 59, 473-494.

Andreoni, J., (1991, "Reasonable Doubt and the Optimal Magnitude of Fines: Should the Penalty Fit the Crime?" Rand Journal of Economics 22(3), 385-95.

Banks, Jeffrey., and J. Sobel, (1987), "Equilibrium Selection in Signaling Games," *Econometrica* 55, 647-661.

Barclay, M., and C. Smith, JR., (1995), "The Maturity Structure of Corporate Debt," *Journal of Finance* 50, 609-631.

Baron, D., and D. Besanko, (1984), "Regulation, Asymmetric Information and Auditing", *Rand Journal of Economics*, 15(4),. 447-470.

Baron,D., and R. Myerson, (1982), "Regulating a Monopolist with Unknown Costs", *Econometrica* 50,911-930

Becker, G.S., (1968), "Crime and Punishment: An Economic Approach", *Journal of Political Economy* 76, 169-217.

Becker, G., (1983), "A Theory of Competition Among Pressure Groups for Political Influence", *Quarterly Journal of Economics* 98, 371-400.

Becker, G.S., and G.J. Stigler, (1974), "Law Enforcement, Malfeasance, and the Compensation of Enforcers", *Journal of Legal Studies 3*, 1-18.

Berlin, M., and J. Loeys, (1988), "Bond Covenants and Delegated Monitoring", *Journal of Finance* 43, 397-412.

Bhattacharya, S., and P, Pfleiderer, (1985), "Delegated Portfolio Management", *Journal of Economic Theory* 36, 1-25.

Billett, M., M. Flannery, and J. Garfinkel, (1995), "The Effect of Lender Identity on a Borrowing Firm's Equity Return", *Journal of Finance* 50, 699-718.

Border, K. C. and J. Sobel, (1987), "Samurai Accounting: A Theory of Auditing and Plunder", Rev. Econ. Studies 54, 525-40.

Campbell, T., and W. Kracaw, (1980), "Information Production, Market Signaling and The Theory of Financial Intermediation," *Journal of Finance* 35, 863-882.

Dewatripont, M., and J. Tirole, (1995), "Advocates", Université de Toulouse Working Paper.

Diamond,D., (1984), "Financial Intermediation and Delegated Monitoring," *Review of Economic Studies* 51, 393-414.

Diamond, D., (1991), "Monitoring and Reputation: The Choice Between Bank Loans and Directly Placed Debt," *Journal of Political Economy* 99, 688-721.

Diamond, D., (1993), "Seniority and Maturity of Debt Contracts," *Journal of Financial Economics* 33, 341-368.

Fama, E., (1985), "What is Different About Banks?" *Journal of Monetary Economics* 15, 29-39.

Gertner, R., and D. Scharfstein, (1991), "A Theory of Workouts and The Effects of Reorganization Law," *Journal of Finance* 46, 1189-1222.

Gertner, R., R. Gibbons and D. Scharfstein, (1988), "Simultaneous Signaling to The Capital and Product Market," *Rand Journal of Economics* 19, 173-190.

Graetz, M., (1986). " The Tax Compliance Game: Toward an Interactive Theory of Law Enforcement." *Journal of Law Economics and Organization* 2, 1-322.

Grossman, S., and O, Hart, (1983), "An Analysis of the Principal-Agent Problem", *Econometrica* 51, 7-46.

Grossman, S., and O.Hart, (1986), "The Cost and Benefits of Ownership: A Theory of Vertical and Lateral Integration," *Journal of Political Economy* 94, 691-719.

Harrington, W., (1988). " Enforcement Leverage when Penalties are Restricted." *Journal of Public Economics* 37, 29-53.

Hart, O., and J, Moore, (1989), "Default and Renegotiation," London School of Business Working Paper.

Hart, O., and J, Moore, (1994), "A Theory of Debt Based on The Inalienability of Human Capital," *Quarterly Journal of Economics* 86, 841-879.

Houston, J., and C. James, (1995), "Information Monopoly and the Mix of Private and Public Debt Claims," University of Florida Working Paper.

James, C., (1987), "Some Evidence on the Uniqueness of Bank Loans," *Journal of Financial Economics* 19, 217-235.

Jensen, M., (1986), "Agency Costs of Free Cash Flow, Corporate Finance and Takeovers," *American Economic Review* 76, 323-329.

Kaplow, L. and S. Shavell (1994). "Optimal Law Enforcement with Self-Reporting of Behavior," *Journal of Political Economy* 102, 583-605.

Kilhstrom, R., (1986), "Optimal Contracts for Security Analysts and Portfolio Managers", Wharton School Working Paper.

Laffont, J., and J, Tirole (1986), "Using Cost Observation to Regulate Firms", *Journal of Political Economy* 94,614-641.

Laffont, J.J. and J. Tirole (1990). "The Politics of Government Decision Making: Regulatory Institutions", *Journal of Law, Economics, and Organization* 6, 1-32.

Laffont, J.J., and J. Tirole (1991). "The Politics of Government Decision Making: A Theory of Regulatory Capture", *Quarterly Journal of Economics* 106, 1089-1127.

Landes, W., and R. Posner (1975). "The Private Enforcement of the Law", *Journal of Legal Studies* 1, 1-46.

Lummer, S., and J. McConnell, (1989), "Further Evidence on the Bank Lending Process and the Capital Market Responses to Bank Loan Agreement", *Journal of Financial Economics* 25, 99-122.

Malik, A., (1990), "Avoidance, Screening and Optimum Enforcement," *Rand Journal of Economics* 21, 341-353.

Mookherjee, D., P., I, (1994). "Marginal Deterrence in Enforcement of Law." *Journal of Political Economy* 102, 1039-1066.

Mookherjee, D., and I. Png (1992). "Monitoring vis-a-vis Investigation in Enforcement of Law," *American Economic Review* 83, 556-65.

Myerson, R., (1979), " Incentive Compatibility and the Bargaining Problem", *Econometrica* 47, 61-74

Park, C., (1994), "Monitoring and Debt Seniority Structure," University of Chicago Working Paper.

Polinsky, M., (1980). "Private vs. Public Enforcement of Fines," *Journal of Legal Studies* 9, 105-27.

Polinsky, M., and S. Shavell (1979). "The Optimal Tradeoff between the Probability and Magnitude of Fine,." *American Economic Review* 880-91.

Raja, R., (1992), " Insiders and Outsiders: The Choice Between Informed and Arm's-Length Debt," *Journal of Finance* 47, 1367-1400.

Rajan,R., and A. Winton, (1995), "Covenants and Collateral as Incentives to Monitor," *Journal of Finance* 50, 1113-1146.

Rothschild, M., and J. Stiglitz, (1970), " Increasing Risk: I. A Definition", *Journal of Economic Theory* 2, 225-243

Stigler, George J., (1970), "The Optimum Enforcement of Laws." *Journal of Political Economy* 78, 526-36.

Stoughton, N., (1993), " Moral Hazard and Portfolio Management Problem", *Journal of Finance* 48, 2009-2028

Tirole, J., (1986), "Hierarchies and Bureaucracies: on the Role of Collusion in Organization", *Journal of Law, Economics, and Organization* 2, 181-214.

Tirole, J., (1992), "Collusion and the Theory of Organizations." *Document De Travail,* 1-61.

BIOGRAPHICAL SKETCH

Wei-Lin Liu received a bachelor's degree in physics in 1989 from Flinders University of South Australia, and a master's degree in physics in 1991 from Virginia Tech. He entered the graduate program in finance in 1993.

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy

David T. Brown

David T. Brown, Chair
Associate Professor of Finance,
Insurance, and Real Estate

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy

Tracy R. Lewis, Cochair
James W. Walter Eminent Scholar
in Entrepreneurship and Economics

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy

Joel T. Houston

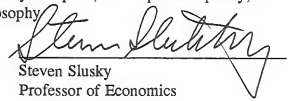
Joel Houston
Associate Professor of Finance,
Insurance, and Real Estate

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy

Michael D. Ryngaert

Michael Ryngaert
Associate Professor of Finance,
Insurance, and Real Estate

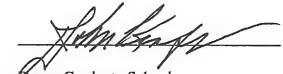
I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.



Steven Slusky
Professor of Economics

This dissertation was submitted to the Graduate Faculty of the Department of Finance, Insurance, and Real Estate in the College of Business Administration and the Graduate School and was accepted as partial fulfillment of the requirements for the degree of Doctor of Philosophy.

August 1997



Dean, Graduate School